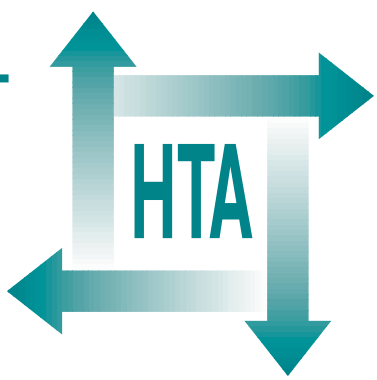
Health Technology Assessment 1998; Vol. 2: No. 12

Review

Postoperative analgesia and vomiting, with special reference to day-case surgery: a systematic review

Henry J McQuay
R Andrew Moore

Health Technology Assessment
NHS R&D HTA Programme



Standing Group on Health Technology

Chair: Professor Sir Miles Irving,
Professor of Surgery, University of Manchester, Hope Hospital, Salford †

Dr Sheila Adam,
Department of Health

Professor Martin Buxton,
Professor of Economics, Brunel University †

Professor Angela Coulter,
Director, King's Fund, London

Professor Anthony Culyer,
Deputy Vice-Chancellor, University of York

Dr Peter Doyle,
Executive Director, Zeneca Ltd,
ACOST Committee on Medical Research & Health

Professor John Farndon,
Professor of Surgery, University of Bristol †

Professor Charles Florey,
Department of Epidemiology & Public Health, Ninewells Hospital & Medical School, University of Dundee †

Professor John Gabbay,
Director, Wessex Institute for Health Research & Development †

Professor Sir John Grimley Evans,
Department of Geriatric Medicine, Radcliffe Infirmary, Oxford †

Dr Tony Hope,
The Medical School, University of Oxford †

Professor Howard Glennester,
Professor of Social Science & Administration, London School of Economics & Political Science

Mr John H James,
Chief Executive, Kensington, Chelsea & Westminster Health Authority

Professor Richard Lilford,
Regional Director, R&D, West Midlands †

Professor Michael Maisey,
Professor of Radiological Sciences, UMDS, London

Dr Jeremy Metters,
Deputy Chief Medical Officer, Department of Health †

Mrs Gloria Oates,
Chief Executive, Oldham NHS Trust

Dr George Poste,
Chief Science & Technology Officer, SmithKline Beecham †

Professor Michael Rawlins,
Wolfson Unit of Clinical Pharmacology, University of Newcastle-upon-Tyne

Professor Martin Roland,
Professor of General Practice, University of Manchester

Mr Hugh Ross,
Chief Executive, The United Bristol Healthcare NHS Trust †

Professor Ian Russell,
Department of Health, Sciences & Clinical Evaluation, University of York

Professor Trevor Sheldon,
Director, NHS Centre for Reviews & Dissemination, University of York †

Professor Mike Smith,
Director, The Research School of Medicine, University of Leeds †

Dr Charles Swan,
Consultant Gastroenterologist, North Staffordshire Royal Infirmary

Dr John Tripp,
Department of Child Health, Royal Devon & Exeter Healthcare NHS Trust †

Professor Tom Walley,
Department of Pharmacological Therapeutics, University of Liverpool †

Dr Julie Woodin,
Chief Executive, Nottingham Health Authority †

† Current members

HTA Commissioning Board

Chair: Professor Charles Florey, Department of Epidemiology & Public Health,
Ninewells Hospital & Medical School, University of Dundee †

Professor Ian Russell,
Department of Health, Sciences & Clinical Evaluation, University of York *

Dr Doug Altman,
Director of ICRF/NHS Centre for Statistics in Medicine, Oxford †

Mr Peter Bower,
Independent Health Advisor, Newcastle-upon-Tyne †

Ms Christine Clark,
Honorary Research Pharmacist, Hope Hospital, Salford †

Professor David Cohen,
Professor of Health Economics, University of Glamorgan

Mr Barrie Dowdeswell,
Chief Executive, Royal Victoria Infirmary, Newcastle-upon-Tyne

Professor Martin Eccles,
Professor of Clinical Effectiveness, University of Newcastle-upon-Tyne †

Dr Mike Gill,
Director of Public Health and Health Policy, Brent & Harrow Health Authority †

Dr Jenny Hewison,
Senior Lecturer, Department of Psychology, University of Leeds †

Dr Michael Horlington,
Head of Corporate Licensing, Smith & Nephew Group Research Centre

Professor Sir Miles Irving
(Programme Director), Professor of Surgery, University of Manchester, Hope Hospital, Salford †

Professor Alison Kitson,
Director, Royal College of Nursing Institute †

Professor Martin Knapp,
Director, Personal Social Services Research Unit, London School of Economics & Political Science

Dr Donna Lamping,
Senior Lecturer, Department of Public Health, London School of Hygiene & Tropical Medicine †

Professor Theresa Marteau,
Director, Psychology & Genetics Research Group, UMDS, London

Professor Alan Maynard,
Professor of Economics, University of York †

Professor Sally McIntyre,
MRC Medical Sociology Unit, Glasgow

Professor Jon Nicholl,
Director, Medical Care Research Unit, University of Sheffield †

Professor Gillian Parker,
Nuffield Professor of Community Care, University of Leicester †

Dr Tim Peters,
Reader in Medical Statistics, Department of Social Medicine, University of Bristol †

Professor David Sackett,
Centre for Evidence Based Medicine, Oxford

Professor Martin Severs,
Professor in Elderly Health Care, Portsmouth University †

Dr David Spiegelhalter,
MRC Biostatistics Unit, Institute of Public Health, Cambridge

Dr Ala Szczepura,
Director, Centre for Health Services Studies, University of Warwick †

Professor Graham Watt,
Department of General Practice, Woodside Health Centre, Glasgow †

Professor David Williams,
Department of Clinical Engineering, University of Liverpool

Dr Mark Williams,
Public Health Physician, Bristol

Dr Jeremy Wyatt,
Senior Fellow, Health and Public Policy, School of Public Policy, University College, London †

* Previous Chair
† Current members

How to obtain copies of this and other HTA Programme reports.
An electronic version of this publication, in Adobe Acrobat format, is available for downloading free of charge for personal use from the HTA website (http://www.hta.ac.uk). A fully searchable CD-ROM is also available (see below).

Printed copies of HTA monographs cost £20 each (post and packing free in the UK) to both public **and** private sector purchasers from our Despatch Agents.

Non-UK purchasers will have to pay a small fee for post and packing. For European countries the cost is £2 per monograph and for the rest of the world £3 per monograph.

You can order HTA monographs from our Despatch Agents:

– fax (with **credit card** or **official purchase order**)
– post (with **credit card** or **official purchase order** or **cheque**)
– phone during office hours (**credit card** only).

Additionally the HTA website allows you **either** to pay securely by credit card **or** to print out your order and then post or fax it.

Contact details are as follows:

HTA Despatch
c/o Direct Mail Works Ltd
4 Oakwood Business Centre
Downley, HAVANT PO9 2NP, UK

Email: orders@hta.ac.uk
Tel: 02392 492 000
Fax: 02392 478 555
Fax from outside the UK: +44 2392 478 555

NHS libraries can subscribe free of charge. Public libraries can subscribe at a very reduced cost of £100 for each volume (normally comprising 30–40 titles). The commercial subscription rate is £300 per volume. Please see our website for details. Subscriptions can only be purchased for the current or forthcoming volume.

Payment methods

Paying by cheque
If you pay by cheque, the cheque must be in **pounds sterling**, made payable to *Direct Mail Works Ltd* and drawn on a bank with a UK address.

Paying by credit card
The following cards are accepted by phone, fax, post or via the website ordering pages: Delta, Eurocard, Mastercard, Solo, Switch and Visa. We advise against sending credit card details in a plain email.

Paying by official purchase order
You can post or fax these, but they must be from public bodies (i.e. NHS or universities) within the UK. We cannot at present accept purchase orders from commercial companies or from outside the UK.

How do I get a copy of *HTA on CD*?

Please use the form on the HTA website (www.hta.ac.uk/htacd.htm). Or contact Direct Mail Works (see contact details above) by email, post, fax or phone. *HTA on CD* is currently free of charge worldwide.

The website also provides information about the HTA Programme and lists the membership of the various committees.

Postoperative analgesia and vomiting, with special reference to day-case surgery: a systematic review

Henry J McQuay
R Andrew Moore

Pain Research & Nuffield Department of Anaesthetics,
University of Oxford,
Oxford Radcliffe Hospital Trust,
Churchill Hospital, Oxford

Published December 1998

This report should be referenced as follows:

McQuay HJ, Moore RA. Postoperative analgesia and vomiting, with special reference to day-case surgery: a systematic review. *Health Technol Assessment* 1998; **2**(12).

Health Technology Assessment is indexed in *Index Medicus*/MEDLINE and *Excerpta Medica*/EMBASE. Copies of the Executive Summaries are available from the NCCHTA web site (see overleaf).

NHS R&D HTA Programme

The overall aim of the NHS R&D Health Technology Assessment (HTA) programme is to ensure that high-quality research information on the costs, effectiveness and broader impact of health technologies is produced in the most efficient way for those who use, manage and work in the NHS. Research is undertaken in those areas where the evidence will lead to the greatest benefits to patients, either through improved patient outcomes or the most efficient use of NHS resources.

The Standing Group on Health Technology advises on national priorities for health technology assessment. Six advisory panels assist the Standing Group in identifying and prioritising projects. These priorities are then considered by the HTA Commissioning Board supported by the National Coordinating Centre for HTA (NCCHTA).

This report is one of a series covering acute care, diagnostics and imaging, methodology, pharmaceuticals, population screening, and primary and community care. It was identified as a priority by the Pharmaceutical Panel and funded as project number 94/11/04.

The views expressed in this publication are those of the authors and not necessarily those of the Standing Group, the Commissioning Board, the Panel members or the Department of Health. The editors wish to emphasise that funding and publication of this research by the NHS should not be taken as implicit support for the recommendations for policy contained herein. In particular, policy options in the area of screening will, in England, be considered by the National Screening Committee. This Committee, chaired by the Chief Medical Officer, will take into account the views expressed here, further available evidence and other relevant considerations.

Reviews in *Health Technology Assessment* are termed 'systematic' when the account of the search, appraisal and synthesis methods (to minimise biases and random errors) would, in theory, permit the replication of the review by others.

The editors have tried to ensure the accuracy of this report but cannot accept responsibility for any errors or omissions. They would like to thank the referees for their constructive comments on the draft document.

ISSN 1366-5278

Published by Core Research, Alton, on behalf of the NCCHTA.
Printed on acid-free paper in the UK by The Basingstoke Press, Basingstoke.

Copies of this report can be obtained from:

The National Coordinating Centre for Health Technology Assessment,
Mailpoint 728, Boldrewood,
University of Southampton,
Southampton, SO16 7PX, UK.
Fax: +44 (0) 1703 595 639 Email: hta@soton.ac.uk
http://www.soton.ac.uk/~hta

Contents

List of abbreviations

AUC	area under a curve*
CATPR	categorical verbal rating scale of pain relief
CER	control event rate
CI	confidence interval
EER	experimental event rate
ESR	erythrocyte sedimentation rate
NNH	number-needed-to-harm
NNT	number-needed-to-treat
NSAID	non-steroidal anti-inflammatory drug
NSD	no significant difference*
PCA	patient-controlled analgesia
PI	pain intensity*
PID	pain intensity difference*
PO	postoperative*
PONV	postoperative nausea and vomiting
PR	pain relief*
RB	relative benefit*
RCT	randomised controlled trial
RR	relative risk*
RWJ	Robert Wood Johnson Pharmaceutical Research Institute, Spring House, PA, USA
SD	standard deviation*
SEM	standard error of the mean*
SPID	sum of pain intensity differences
TENS	transcutaneous electrical nerve stimulation
TFA	time to first analgesic*
TOTPAR	total pain relief
VAS	visual analogue scale
VRS	verbal rating scale*

* Used only in figures and tables

Executive summary

Background

Day-case surgery is of great value to patients and the health service. It enables many more patients to be treated properly, and faster than before. Newer, less invasive, operative techniques will allow many more procedures to be carried out.

There are many elements to successful day-case surgery. Two key components are the effectiveness of the control of pain after the operation, and the effectiveness of measures to minimise postoperative nausea and vomiting.

Objectives

To enable those caring for patients undergoing day-case surgery to make the best choices for their patients and the health service, this review sought the highest quality evidence on:

- the effectiveness of the control of pain after an operation
- the effectiveness of measures to minimise postoperative nausea and vomiting.

Methods

Full details of the search strategy are presented in the report.

Results

Analgesia

The systematic reviews of the literature explored whether different interventions work and, if they do work, how well they work. A number of conclusions can be drawn.

Ineffective interventions

There is good evidence that some interventions are ineffective. They include:

- transcutaneous electrical nerve stimulation in acute postoperative pain
- the use of local injections of opioids at sites other than the knee joint
- the use of dihydrocodeine, 30 mg, in acute postoperative pain (it is no better than placebo).

Interventions of doubtful value

Some interventions may be effective but the size of the effect or the complication of undertaking them confers no measurable benefit over conventional methods. Such interventions include:

- injecting morphine into the knee joint after surgery: there is a small analgesic benefit which may last for up to 24 hours but there is no clear evidence that the size of the benefit is of any clinical value
- manoeuvres to try and anticipate pain by using pre-emptive analgesia; these are no more effective than standard methods
- administering non-steroidal anti-inflammatory drugs (NSAIDs) by injection or per rectum in patients who can swallow; this appears to be no more effective than giving NSAIDs by mouth and, indeed, may do more harm than good
- administering codeine in single doses; this has poor analgesic efficacy.

Interventions of proven value

These include a number of oral analgesics including (at standard doses):

- dextropropoxyphene
- tramadol
- paracetamol
- ibuprofen
- diclofenac.

Diclofenac and ibuprofen at standard doses give analgesia equivalent to that obtained with 10 mg of intramuscular morphine. Each will provide at least 50% pain relief from a single oral dose in patients with moderate or severe postoperative pain. Paracetamol and codeine combinations also appear to be highly effective, although there is little information on the standard doses used in the UK. The relative effectiveness of these analgesics is compared in an effectiveness 'ladder' which can inform prescribers making choices for individual patients, or planning day-case surgery. Dose–response relationships show that higher doses of ibuprofen may be particularly effective.

Topical NSAIDs (applied to the skin) are effective in minor injuries and chronic pain but there is no obvious role for them in day-case surgery.

Postoperative nausea and vomiting

The proportion of patients who may feel nauseated or vomit after surgery is very variable, despite similar operations and anaesthetic techniques. Systematic review can still lead to clear estimations of effectiveness of interventions. Whichever anti-emetic is used, the choice is often between prophylactic use (trying to prevent anyone vomiting) and treating those people who do feel nauseated or who may vomit.

Systematic reviews of a number of different anti-emetics show clearly that none of the anti-emetics is sufficiently effective to be used for prophylaxis. Moreover, a cost-effectiveness analysis shows that prophylaxis, especially with newer anti-emetics, not only does not prevent any more people from vomiting or feeling nauseated than treating established nausea or vomiting, but exposes patients to considerably more drug at considerably higher cost.

Conclusions

This report has focused on two elements of day-case care. It is clear that the economics of day-case work require that the vast majority of patients are fit to go home and that, once home, they do not have to return to hospital or seek advice from the primary care team. To date, audits at local level have identified both pain and nausea and vomiting as problems. Providing adequate analgesia may be easier than guaranteeing minimal nausea and vomiting. The package of care in day-case surgery needs to be revisited regularly lest surgical interventions are the cause of increased hospitalisation and returns.

Research recommendations

- To extend the number of systematic reviews to include other analgesics, including the newer NSAIDs. This would provide a more comprehensive ladder of relative efficacy. It is unlikely that large 'head-to-head' comparisons of analgesics in randomised controlled trials would provide more useful information.

- To establish pilot audits of the implementation of the information included in this report; both before and after audits are needed to put the existing evidence into clinical practice to good effect.

- To investigate the effect of randomness in clinical trials. Because there are substantial numbers of analgesic trials and they are usually performed using standard methods and including patients with similar entry criteria (moderate or severe pain), they may be usefully studied to examine the effects of randomness in clinical trials. Variability between trials is large and understanding the effects of chance would help to inform us of how large trials need to be to give an accurate clinical feel for a new drug.

- To investigate how to minimise postoperative nausea and vomiting. This varies considerably between trials and may be the result of random chance, but it is just as likely that components of the overall package of care other than anaesthesia or anti-emetics are important. There is an obvious and important research agenda here in understanding how best to minimise postoperative nausea and vomiting. However, this is a complex area which will not easily be understood.

Chapter 1
Introduction

A simple way of looking at day-case surgery is that both pain relief and postoperative nausea and vomiting (PONV) should be well controlled. Patients should neither have to stay in hospital (because of poor pain relief or PONV) nor have to contact healthcare professionals after they have left hospital.

We have therefore concentrated on developing two league tables, or rank ordering. One is for which pain-killer or analgesic works best by mouth after surgery. This has produced surprises – the 'standard' take-home analgesic package often contains analgesics which do poorly in the ranking. Also, the efficacy of oral pain-killers has been compared with injections of morphine or non-steroidal anti-inflammatory drugs (NSAIDs). The net has been thrown quite wide and information on transcutaneous electrical nerve stimulation (TENS) and injections of morphine into the knee joint and other peripheral sites has been included, together with information on whether the timing of the analgesic (before or after surgery) makes any clinical difference.

Producing this ranking has involved developing new methods and extending existing ones. The first part of the report documents the methods because, although the required information is that in the league-table, the credibility of the league table rests on the credibility of the methods used to derive it.

Similar methods have been used to tackle treatment of PONV. A preliminary league-table of various prophylactic measures is presented in the report. None of them appear to work particularly well. There is much less information on treatment (rather than prevention) of PONV but some useful results are presented.

It is hard to be precise about the effect that poorly controlled pain or poorly controlled PONV have on the incidence of patients having to stay in hospital, or on the incidence of consultations after leaving hospital. Ideal targets seem to be that less than 1% of patients should have to stay in, and less than 1% should have to consult.[1] Audits show that poor pain control can certainly result in rates greater than 1% for both categories,[2] and that providing better pain control produced a worthwhile reduction in both types of problem.

The authors hope that this report will enable a higher quality of evidence to be used for future guidelines, for pain relief both in and out of hospital, and for prevention of PONV.

Background and key questions

The rational approach to postoperative care is to use the highest quality evidence available and, in this context, this comes from systematic reviews of valid randomised trials. The results will still have to be adapted to the circumstance of the individual but our chosen policies will be more discerning. Treatments which are simple, clinically appropriate and evidence-based are focused on in this study. The authors were fortunate that there is now a steady supply of systematic reviews available in the pain world[3] (an updated listing is available on the Internet at <http://www.jr2.ox.ac.uk/Bandolier/painres/MApain.html>).

To achieve the best outcome it has been necessary to pick our way through the evidence. There is a complicated relationship between evidence, guidelines, research and legal considerations, and the patients' outcomes as assessed by audit (see *Figure 1*).

High-quality postoperative care needs effective pain management. While we would like to believe that we practice good pain control, a survey[4] of

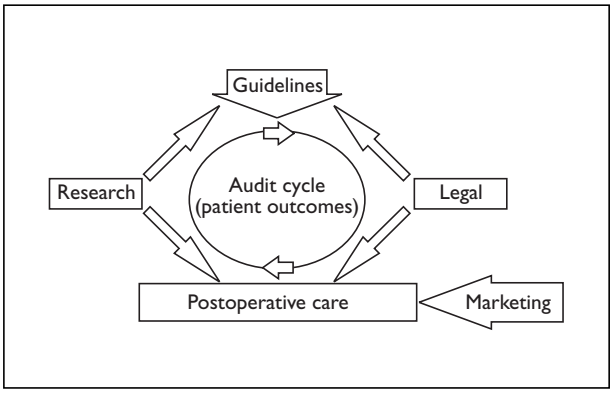


FIGURE 1 *Influences on postoperative care*

TABLE 1 *Inpatient survey*[4]

	Number of patients	Percentage
Pain was present all or most of the time	1042/3162	33
Pain was severe or moderate	2755/3157	87
Pain was worse than expected	182/1051	17
Had to ask for drugs	1085/2589	42
Drugs did not arrive immediately	455/1085	41

5150 recently discharged patients from 36 UK hospitals showed that, in the 3163 who responded to questions on pain, practice was far from ideal (*Table 1*). The questions asked are useful for audit.

Postoperative care is more than a collection of interventions. It is a package of care that needs to be examined as a whole as well as in its parts. Publications which analyse the process of post-operative care provision are rare, perhaps because they attract few academic plaudits. There is good evidence that the risk of adverse events is increased when high-technology approaches are used for drug administration.[5]

Pain charts

Pain charts used as part of normal practice will improve quality of care.[6,7] The fact of a chart is more important than its form, with pain measurements recorded at the same time as sedation, respiratory frequency and nausea, and as part of ongoing audit. An example is the Burford chart.[8] There are special scales for children.[9]

Acute pain services

One remedy for poor management is the provision of an acute pain service.[10] The dispute about what should be provided ranges from a full 'menu', including all the high-technology options,[11] to a service limited to supervision of good practice guidelines for low-technology approaches and staff education.[6,7] Training and education should be the main tasks of an acute pain service.

Key questions

Several key questions for pain relief are obvious.

- The simplest observation is that if patients can swallow then the oral route should be preferred.
- The next relates to prophylaxis versus treat-as-necessary: are the arguments for prophylaxis convincing?
- A third is whether or not injections of local anaesthetic, with or without opioids, are useful in the day surgery setting.

For the future, we need to know why, in some patients, acute pain becomes chronic pain.

Can the patient swallow?

Most postoperative pain is managed solely with medication. Perhaps because anaesthetists work with injected drugs, there is a natural belief that drugs which are injected are more powerful than drugs taken by mouth.

Some of the questions which need to be answered include:

- which classes of drugs are the most effective postoperative analgesics? (or which are least effective?)
- within a class of drugs, does the same dose work better when injected or when taken orally?

It is important to know which oral analgesics to recommend to patients because so much post-operative care is now in the home. It is biased to think of patients after major surgery but they too need oral analgesics when they can swallow. The evidence from trials in which drugs are compared with placebo may be used to build a ranking of relative efficacy.

For this study, all the trials of a particular drug compared with placebo in postoperative pain were obtained. The drug's performance in the trials was then converted into a common currency, viz. the proportion of patients with moderate or severe postoperative pain who achieved at least 50% pain relief compared with placebo over 6 hours.

The most effective drugs have a low number-needed-to-treat (NNT) of about 2, meaning that for every two patients who receive the drug one will achieve at least 50% relief because of the treatment (the other patient may obtain relief but it does not reach the 50% level). The NNT is treatment-specific, which is useful for comparison of relative efficacy but, because these NNT comparisons are against placebo, the best NNT of 2 means that while 50 of 100 patients will get at least 50% relief because of the treatment, another 20 patients will have a placebo response which gives them at least 50% relief; hence, with ibuprofen 70 patients from 100 will have effective pain relief.

For paracetamol, 1 g, the NNT is 4 (see chapter 8). Combination of paracetamol with codeine, 60 mg, improves the NNT to 3. Ibuprofen is better at 2. The clear message is that, of the oral analgesics, NSAIDs perform best and that paracetamol alone or in combination is also effective. The strongest oral analgesic regimen would be an oral NSAID supplemented as necessary with paracetamol and opioid. As pain wanes, then the prescription should be paracetamol-based, supplemented if necessary by an NSAID. When used in day surgery, a regimen like this results in high-quality pain relief without recourse to general practitioner visits.[2]

Even if patients can swallow, is it best to give drugs by injection or suppository?

There is no evidence that NSAIDs given rectally or by injection perform any better than the same drug at the same dose given by mouth (see chapter 10); two randomised, double-blind placebo-controlled comparisons of oral ibuprofen arginine, 400 mg, failed to distinguish any difference from intramuscular ketorolac, 30 mg.[12,13] These other non-oral routes become appropriate when patients cannot swallow.

The patient can't swallow analgesics

There is, as yet, no information on the relative efficacy of injected opioids or NSAIDs and regional anaesthetic techniques. However, some statements can be made:

- injecting morphine at a dose of 10 mg provides similar analgesia to oral NSAID[14]
- injecting morphine at doses of 10–20 mg provides similar analgesia to injected NSAID[15]
- injecting NSAID provides similar analgesia to oral NSAID[12,13]
- injecting 20 mg of morphine provides greater analgesia than injecting 10 mg, and greater analgesia than the best performers in the oral league table.[15]

Other techniques

TENS

TENS is not effective in postoperative pain[16,17] (see chapter 19) and is of limited value in labour pain[18] (see chapter 20).

Psychological methods

There is evidence that psychological approaches are beneficial.[19] Cognitive behavioural methods can reduce pain and distress in patients with burns. Preparation before surgery can reduce postoperative analgesic consumption. The evidence for the use of relaxation and music on postoperative pain is confounded by the poor quality of trials.[20]

Pain which persists: prophylaxis or wait till it happens?

The intriguing questions in acute pain are the following.

- Is there a link between bad, acute postoperative pain and perseveration of this pain into a chronic status?
- Can anything be done to prevent this?

What remains unexplained is why some patients end up with chronic pain after surgery when others do not. A simplistic explanation is that the chronic pain results from nerve damage at surgery. An alternative explanation is that it is those patients with severe postoperative pain who develop the chronic pain. The easy linkage is then to propose that if the acute pain was better controlled then chronic pain would not develop.[21,22]

Pre-emptive short-term analgesia

The evidence for a clinical advantage in giving an intervention before pain as opposed to giving the same intervention after pain is still unconvincing[23] (see chapter 18). Certainly by far the majority of trials of pre-pain versus post-pain medication has failed to show any clinically meaningful benefit.

Peripheral opioids

For intra-articular peripheral opioids at least, the story becomes a little clearer. A systematic review of valid trials of intra-articular morphine in knee surgery has shown that morphine in the knee joint can indeed provide analgesia[24] (see chapter 17). This analgesia can continue for up to 24 hours, although there is no dose response. It is the long duration of action which suggests this technique might have practical application beyond its research interest.[25]

Postoperative nausea and vomiting

Managing PONV well is part of quality post-operative care. Recent evidence from systematic reviews of PONV shows that prophylactic anti-emetics are less effective than might have been hoped (see chapter 20). Use of propofol,[26] omitting nitrous oxide,[27] and different anaesthetic techniques[28] all have similar efficacy, preventing about one patient vomiting for every four or five treated. This is also the case in high-risk settings such as strabismus surgery.[29] There is much less data for treatment of established PONV than for prophylaxis but treatment with ondansetron appears to have a similar level of efficacy to prevention.[30]

Conclusion

The availability of high-quality evidence gives a firm foundation for building better postoperative care. It enables informed decisions to be made about drugs and route of administration for individual patients and services, both for pain and for nausea and vomiting. Bringing this together into an efficient and effective service will be the challenge, so that audit or controlled trials can demonstrate the effectiveness of our postoperative care.

References

1. Millar JM, Rudkin GE, Hitchcock M. Practical anaesthesia and analgesia for day surgery. Oxford: Bios, 1997.

2. Haynes TK, Evans DEN, Roberts D. Pain relief after day surgery: quality improvement by audit. *J 1-day Surg* 1995;(Summer):12–15.

3. Jadad AR, McQuay HJ. Meta-analyses to evaluate analgesic interventions: a systematic qualitative review of their methodology. *J Clin Epidemiol* 1996;49:235–43.

4. Bruster S, Jarman B, Bosanquet N, Weston D, Erens R, Delbanco TL. National survey of hospital patients. *BMJ* 1994;**309**:1542–6.

5. Bates DW, Cullen DJ, Laird N, Petersen LA, Small SD, Servi D, *et al.* Incidence of adverse drug events and potential adverse drug events. *JAMA* 1995;**274**:29–34.

6. Gould TH, Crosby DL, Harmer M, Lloyd SM, Lunn JN, Rees GAD, *et al.* Policy for controlling pain after surgery: effect of sequential changes in management. *BMJ* 1992;**305**:1187–93.

7. Rawal N, Berggren L. Organization of acute pain services: a low-cost model. *Pain* 1994;**57**:117–23.

8. Burford Nursing Development Unit. Nurses and pain. *Nurs Times* 1984;**18**:94.

9. McGrath PJ, Ritchie JA, Unruh AM. Paediatric pain. In: Carroll D, Bowsher D, editors. Pain management and nursing care. Oxford: Butterworth Heinemann, 1993: 100–23.

10. Royal College of Surgeons of England and the College of Anaesthetists. Report of the working party on pain after surgery. London: Royal College of Surgeons, 1990.

11. Ready LB, Oden R, Chadwick HS, Benedetti C, Rooke GA, Caplan R, *et al.* Development of an anesthesiology based postoperative pain management service. *Anesthesiology* 1988;**68**:100–6.

12. Laveneziana D, Riva A, Bonazzi M, Cipolla M, Migliavacca S. Comparative efficacy of oral ibuprofen arginine and intramuscular ketorolac in patients with postoperative pain. *Clin Drug Invest* 1996;**11**:8–14.

13. Pagnoni B, Vignali M, Colella S, Monopoli R, Tiengo N. Comparative efficacy of oral ibuprofen arginine and intramuscular ketorolac in patients with postcaesarean section pain. *Clin Drug Invest* 1996;**11**:15–21.

14. Mansfield M, Firth F, Glynn C, Kinsella J. A comparison of ibuprofen arginine with morphine sulphate for pain relief after orthopaedic surgery. *Eur J Anaesthesiol* 1996;**13**:492–7.

15. Norholt S, SindetPedersen S, Larsen U, Bang U, Ingerslev J, Nielsen O, *et al.* Pain control after dental surgery: a double-blind, randomised trial of lornoxicam versus morphine. *Pain* 1996;**67**:335–43.

16. Reeve J, Menon D, Corabian P. Transcutaneous electrical nerve stimulation (TENS): a technology assessment. *Int J Technol Assess Health Care* 1996;**12**:299–324.

17. Carroll D, Tramèr M, McQuay H, Nye B, Moore A. Randomization is important in studies with pain outcomes: systematic review of transcutaneous electrical nerve stimulation in acute postoperative pain. *Br J Anaesth* 1996;**77**:798–803.

18. Carroll D, Tramèr M, McQuay H, Nye B, Moore A. Transcutaneous electrical nerve stimulation in labour pain: a systematic review. *Br J Obstet Gynaecol* 1997;**104**:169–75.

19. Justins DM, Richardson PH. Clinical management of acute pain. *Br Med Bull* 1991;**47**:561–83.

20. Good M. Effect of relaxation and music on postoperative pain: a review. *J Adv Nurs* 1996;**24**:905–14.

21. Eija K, Tiina T, Pertti NJ. Amitriptyline effectively relieves neuropathic pain following treatment of breast cancer. *Pain* 1996;**64**:293–302.

22. Katz J, Jackson M, Kavanagh B, Sandler A. Acute pain after thoracic surgery predicts long-term post-thoracotomy pain. *Clin J Pain* 1996;**12**:50–5.

23. McQuay HJ. Pre-emptive analgesia: a systematic review of clinical studies. *Ann Med* 1995;**27**:249–56.

24. Kalso E, Tramèr M, Carroll D, McQuay H, Moore RA. Pain relief from intra-articular morphine after knee surgery: a qualitative systematic review. *Pain* 1997;**71**:642–51.

25. Kanbak NM, Akpolat N, Öcal T, Doral NM, Ercan M, Erdem K. Intra-articular morphine administration provides pain relief after knee arthroscopy. *Eur J Anaesthesiol* 1997;**14**:153–6.

26. Tramèr M, Moore A, McQuay H. Propofol anaesthesia and postoperative nausea and vomiting: quantitative systematic review of randomized controlled studies. *Br J Anaesth* 1997;**78**:247–55.

27. Tramèr M, Moore A, McQuay H. Omitting nitrous oxide in general anaesthesia: meta-analysis of intraoperative awareness and postoperative emesis in randomized controlled trials. *Br J Anaesth* 1996;**76**:186–93.

28. Tramèr M, Moore A, McQuay H. Meta-analytic comparison of prophylactic anti-emetic efficacy for postoperative nausea and vomiting. Propofol anaesthesia vs. omitting nitrous oxide vs. total i.v. anaesthesia with propofol. *Br J Anaesth* 1997;**78**:256–9.

29. Tramèr M, Moore A, McQuay H. Prevention of vomiting after paediatric strabismus surgery: a systematic review using the numbers-needed-to-treat method. *Br J Anaesth* 1995;**75**:556–61.

30. Tramèr MR, Moore RA, Reynolds DJM, McQuay HJ. A quantitative systematic review of ondansetron in treatment of established postoperative nausea and vomiting. *BMJ* 1997;**314**:1088–92.

Chapter 2

Finding all the relevant trials

Relevant and valid evidence is necessary for effective care. The randomised controlled trial (RCT) is the most reliable way to estimate the effect of an intervention. The principle of randomisation is simple. With randomisation, patients taking part in a randomised trial have the same probability of receiving any of the interventions being compared. Randomisation abolishes selection bias by preventing the investigators influencing allocation of the interventions. It also helps to ensure that other factors, such as age or sex distribution, are equivalent for the different treatment groups. Inadequate randomisation, or inadequate concealment of randomisation, lead to exaggeration of therapeutic effect.[1]

For reviews of evidence to be valid, they need to be systematic; to be systematic, qualitative or quantitative, they need to include all relevant RCTs. Identifying **all** relevant trials is a 'fundamental challenge',[2] which is easily underestimated.

The first obstacle faced by any reviewer is finding out how many eligible RCTs exist. Commonly the total is unknown. Only for newer interventions are reviewers likely to be sure that they have found all relevant RCTs. Otherwise the total number of trials can only be identified by scanning every record in each of the available bibliographic databases, by searching manually all non-indexed journals, theses, proceedings and textbooks, by searching the reference lists of all the reports found, and by asking the investigators of previous RCTs for other published or unpublished information *(Figure 2)*.[3]

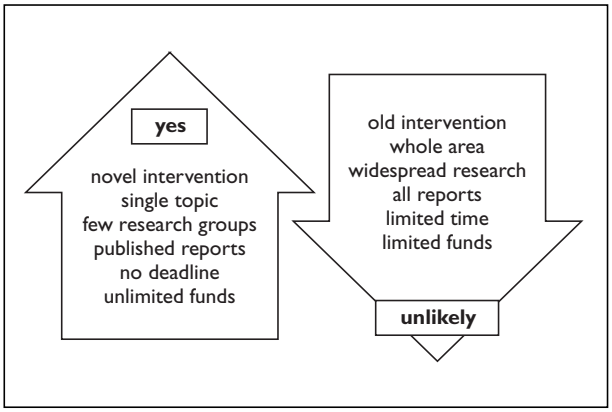


FIGURE 2 *Is a complete search possible?*

In practice, constrained by time and cost, reviewers have to compromise, and hope that what they have found is a representative sample of the unknown total number of eligible trials. The more comprehensive the searching the more trials will be found, and any conclusions will then be stronger. Comprehensive searching can be very time-consuming and costly; again, this emphasises the necessary compromise, where the target is the highest possible yield for given resources.

'Retrieval bias' is the failure to identify reports which could have affected the results of a systematic review or meta-analysis.[4] The failure may be because trials are still ongoing, or completed but as yet unpublished (publication bias) or because although published the search did not find them. Trying to identify unpublished trials by surveying researchers had a very low yield,[5] and is not cheap. Registers of ongoing and completed trials are another means of finding unpublished data but such registers are rare.

In this chapter we describe:

- the methods used to identify eligible reports of RCTs published from 1950 to date
- information management.

Developing a citation database

The process had three phases: definition of inclusion criteria, identification of reports, and information management.

Inclusion criteria

A report was regarded as eligible if both the following criteria were fulfilled.

- Allocation of patients to the intervention was described as randomised (no precise description of the method of randomisation was required), or as double-blind, or as both, or if it was suggested that the interventions were given at random and/or under double-blind conditions.
- Analgesic interventions with pain or adverse effects as outcomes, and/or any intervention using pain as an outcome measure, were compared.

Reports were excluded which investigated analgesic effectiveness during (as opposed to after) diagnostic or surgical procedures.

Identification of reports

Details of the process are presented elsewhere.[6] Since that publication, the major changes are in the use of other databases as well as MEDLINE. Searching EMBASE, the Cochrane Library, CINAHL and PsycLIT is now part of our standard operating procedures (see *Figure 3*).

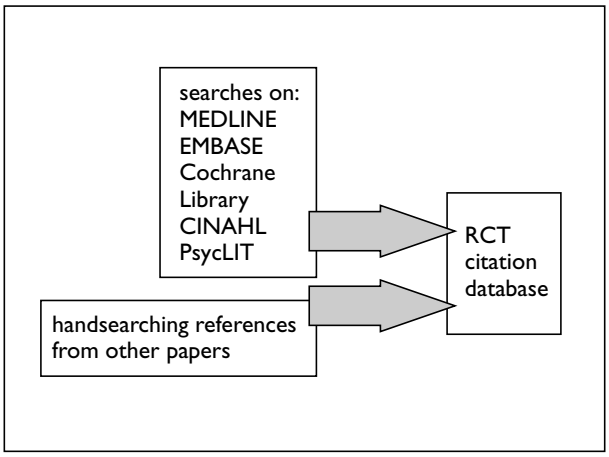


FIGURE 3 *Finding the citations*

MEDLINE search for RCTs published from 1966 to date

The records identified by the optimised MEDLINE search strategy *(Table 2)* were downloaded (Biblio-Link version 1.1, Personal Bibliographic Software, Inc.) and transferred to a reference management program (Pro-Cite, Personal Bibliographic Software, Inc., version 2.1). The records were then sorted in alphabetical order and each downloaded record was checked on-screen for definite eligibility, probable eligibility or ineligibility and coded accordingly within each Pro-Cite record. Hard copies of eligible and probable documents were obtained and, if necessary, translated, and eligibility was then confirmed.

Hand-searching of journals published from 1950 to date

A Pro-Cite file of all the records regarded as eligible and probably eligible (1950–90) was created. This file was used to produce a list of the 50 journals with the highest yield. These journals were then searched by hand to find RCTs. These RCTs, either missed by MEDLINE indexing, or in non-indexed journals, were then added to the citation database if perusal of the hard copy confirmed that they were indeed RCTs.

TABLE 2 *Refined high-yield MEDLINE search strategy*

Step number	Request
1	PAIN*
2	explode PAIN / all subheadings in MeSH
3	ANALG*
4	explode ANALGESIA / all subheadings in MeSH
5	explode ANALGESICS / all subheadings in MeSH
6	CLINICAL
7	TRIALS
8	CLINICAL TRIALS
9	EXPLODE CLINICAL-TRIALS / all subheadings in MeSH
10	RANDOM*
11	RANDOM-ALLOCATION (term allows no subheadings) in MeSH
12	RANDOMIZED-CONTROLLED-TRIALS / all subheadings in MeSH
13	DOUBLE
14	BLIND
15	DOUBLE BLIND
16	DOUBLE-BLIND-METHOD (term allows no subheadings) in MeSH
17	META-ANALYSIS
18	META-ANALYSIS (term allows no subheadings) in MeSH
19	HUMAN
20	(#1 or #2 or #3 or #4 or #5) and (HUMAN in MeSH)
21	HUMAN
22	(#8 or #9 or #10 or #11 or #12 or #15 or #16 or #17 or #18) and (HUMAN in MeSH)
23	#20 and #22

Management of the information

The citation database is maintained as a Pro-Cite file. The number in that database is used as the unique identifying number for the hard copy.

Trends in the numbers of RCTs in pain relief research published, 1950–90

For 1956–80, there were twice as many reports of RCTs published in each successive 5-year period. For 1980–90 the number of reports increased by more than 1000 per 5-year period. More than 85% of the reports identified were published in the last 15 years. This is illustrated by the trend in the number of RCTs published in the journal *Pain* over the past 20 years *(Figure 4)*.

A simple breakdown *(Table 3)* showed that 54% of the reports were in acute pain, 43% in chronic non-cancer and 3% in chronic cancer. Pharmacological reports were commonest (75%), with 14% classified as invasive, 7% as reports of physical interventions, and 2% each for psychological and complementary treatments.

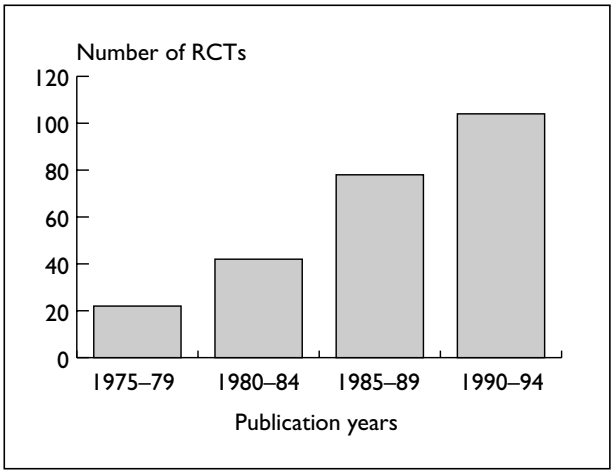


FIGURE 4 *RCTs in the journal* Pain *from 1975 to 1994*

TABLE 3 *Pain trials database 1950–94*

	Acute	Chronic	Cancer	Total	%
Complementary	112	223	10	345	2
Invasive	1697	336	34	2067	14
Pharmacological	5390	4978	337	10,705	75
Physical	402	501	36	939	7
Psychological	100	191	10	301	2
Total	7701	6229	427	14,357	
%	54	43	3		

Conclusions

The importance of basing systematic reviews on the highest quality evidence (randomised trials) is obvious from our experience in the pain field,[7] and from the experience of others. This means that very considerable time and effort has to be spent to gather all the relevant material for each review.

The process described here gives an outline of what is a laborious task. The addition of another year's citations, maintaining the existing database (now 15,000 citations), and the associated chores are a full-time job. To make the information accessible to others we have contributed our citations of known RCTs to the Cochrane Library, and to the compilers of MEDLINE (National Library of Medicine), to ensure that all the RCTs found only by hand-searching are tagged.

References

1. Schulz KF, Chalmers I, Hayes RJ, Altman DG. Failure to conceal treatment allocation schedules in trials influences estimates of treatment effects. *Control Clin Trials* 1994;**15**:63–4S.
2. Chalmers I, Dickersin K, Chalmers T. Getting to grips with Archie Cochrane's agenda. *BMJ* 1992;**305**:786–7.
3. Jadad AR, McQuay HJ. A high-yield strategy to identify randomized controlled trials for systematic reviews. *Online J Curr Clin Trials* 1993; Doc no 33.
4. Simes RJ. Confronting publication bias: a cohort design for meta-analysis. *Stat Med* 1987;**6**:11–29.
5. Hetherington J, Dickerson K, Chalmers I, Meinert C. Retrospective and prospective identification of unpublished controlled trials: lessons from a survey of obstetricians and pediatricians. *Pediatrics* 1989;**84**:374–80.
6. Jadad AR, Carroll D, Moore A, McQuay H. Developing a database of published reports of randomised clinical trials in pain research. *Pain* 1996;**66**:239–46.
7. Carroll D, Tramèr M, McQuay H, Nye B, Moore A. Randomization is important in studies with pain outcomes. Systematic review of transcutaneous electrical nerve stimulation in acute postoperative pain. *Br J Anaesth* 1996;**77**:798–803.

Chapter 3

Judging the quality of the trials

Once all the reports of the trials relevant to your question have been found, there is another stage in the process. This is to confirm that, first, these reports meet certain quality standards and, second, even though a report may pass those quality standards, whether the trial is valid. Imagine a situation where 40 reports of relevant trials were found. Twenty of the reports say that the intervention is terrific, and 20 conclude that it should never be used. Delving deeper, the 20 'negative' reports are found to score highly on your quality standards scale, whereas the 20 'positive' reports score poorly. The quality scale should include measures of bias. Bias is the simplest explanation why poor quality reports give more positive conclusions than high quality reports.

The quality standards which are required cannot be absolute, because for some clinical questions there may not be any RCTs. Setting RCTs as a minimum absolute standard would therefore be inappropriate for some of the questions to which we might want answers. In the study of pain, however, there are two reasons for setting this high standard, and requiring trials to be randomised. The first reason is that there are, particularly for drug interventions, quite a number of RCTs. The second, we would argue, is that that it is even more important to stress the minimum quality standards of randomisation and double-blinding when the outcome measures are subjective.

This chapter describes briefly the development of a quality scale which was then used for the systematic reviews which follow. A detailed description of the way the scale was developed and tested has been published.[1] The chapter concludes with our current views on this and other quality scales.

Developing and validating a quality scale

Previous methods to measure the 'quality' of clinical reports and incorporate the results in systematic reviews may all be criticised because of failure to define quality and because they were not validated.[2–16] The danger is that using these scales might lead to conclusions in the review as inconsistent and unreliable as the component studies.

What makes a trial worthy of the label 'high quality'? Quality could refer to the clinical relevance of the study, to the likelihood of biased results, to the appropriateness of the statistical analysis, to the presentation of the data, or to the ethical implications of the intervention or to the literary style of the manuscript. We consider that quality must primarily indicate the likelihood that the study design reduced bias. Only by avoiding bias is it possible to estimate the effect of a given intervention with any confidence.

The purpose of our scale is to assess the likelihood of the trial design to generate unbiased results and approach the 'therapeutic truth'. This has also been described as 'scientific quality'.[17] Other trial characteristics, such as clinical relevance of the question addressed, data analysis and presentation, literary quality of the report or ethical implications of the study, are not included in our definition.

The aims of the scale are as follows.

1. To assess the scientific quality of any clinical trial in which pain is an outcome measure or in which analgesic interventions are compared for outcomes other than pain (e.g. a study looking at the adverse effect profile of different opioids).
2. To allow consistent and reliable assessment of quality by raters with different backgrounds, including researchers, clinicians, professionals from other disciplines, and members of the general public.

The judges

A multidisciplinary panel of six judges was assembled (a psychologist, a clinical pharmacologist, a biochemist, two anaesthetists, and a research nurse), all with an interest in pain research. The definition of quality and the purposes of the scale were discussed. Each judge then had to produce a list of suggested items to be included on the scale. To generate the items, the judges used both criteria published previously and their own judgement. The suggestions were then combined in a single list of 49 items.

Using a modified nominal group approach to reach consensus,[18] the judges assessed the face

validity of each of the items, according to established criteria.[19] Items associated with low face validity were deleted. An initial instrument was created from the remaining items.

The initial instrument was pre-tested by three raters on 13 study reports and problems in clarity and/or application of each of the items were identified. The panel of judges then modified the wording of the items accordingly and produced detailed instructions describing how each of the items should be assessed and scored. The items were classified by their ability to reduce bias (direct or indirectly), and individual scores were allocated to them by consensus. The frequency of endorsement, consistency and validity of each item were then assessed (see *Figure 5*).

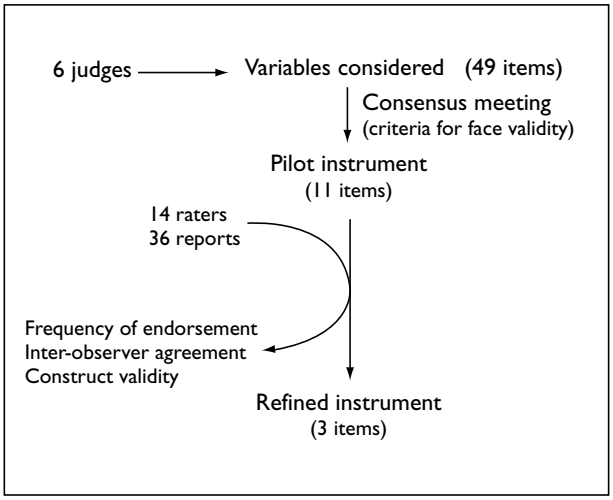


FIGURE 5 *Developing the scale*

Final version of the scale

The final version of the scale contains the three items with highest frequency of endorsement (see *Table 4*). Advice on using the scale is presented in *Table 5,* and the method of scoring RCTs is shown in *Table 6*.

Open versus blind assessments

A chastening finding during the development of the scale was that blind assessment (not knowing authors, journal, year, etc.) of reports produced significantly lower and more consistent quality scores than open assessments.[1] This has important implications, because the cost of organising truly blind assessment is very considerable.

Comments on the scale

The three-point scale is simple, short, valid and reliable. The results suggest that even for those without clinical or research experience in pain relief, it should be possible to score the quality of research reports consistently. Our particular purpose was to allow differential analysis to be undertaken within our systematic reviews based on the quality of the individual primary studies – but the scale may have much wider use.

Chalmers suggested many years ago that the quality of clinical reports should be assessed blind.[3] In our study, such blinded assessment was found to produce significantly lower scores. This may be very important if absolute cut-off scores are imposed by systematic reviewers, and if quality scores are used to weight the results of primary studies in subsequent meta-analysis.[16,20] The results of open evaluations are good enough for busy readers. The improved reliability with blind testing is of more relevance to journal editors, for manuscript selection, and to systematic reviewers. Quality scales without clinimetric evaluation have already been used in pain work to support the conclusions of systematic reviews.[11,13,14]

None of the items on the scale are specific to pain studies. The three items are very similar to the components of a scale used extensively to assess the effectiveness of interventions during pregnancy and childbirth,[8] and also appear in most other scales. Control of selection bias and rater bias is obviously regarded as crucial to quality.

Selection bias is best controlled by allocating patients at random to the different study groups. Each patient should have the same probability of being included in each comparison group, and the allocation should be concealed until after the patient has consented to take part. Methods of allocation based on alternation, date of birth or hospital record number cannot be regarded as random. Failure to secure proper randomisation increases the likelihood that potential participants in a 'randomised' study will be admitted to the study selectively because of prior knowledge of the group to which they would be allocated or excluded selectively before formal admission to the study.[21] Ideal methods of randomisation are those in which individuals with no direct relationship to the study participants are in charge of the allocation (e.g. allocation by telephone from a central coordinating office, concealed from the investigators). Appropriate simpler alternatives are coin tossing, tables of random numbers and numbers generated by computers, but these carry a higher risk of selective selection.

All of these methods are regarded as appropriate for the purposes of our scale, although we are

TABLE 4 *Scale (3-point) to measure the likelihood of bias in pain research reports*

This is not the same as being asked to review a paper. It should not take more than 10 minutes to score a report and there are no right or wrong answers.	
Please read the article and try and answer the following questions (see attached instructions).	
1. Was the study described as randomised (this includes the use of words such as randomly, random and randomisation)? 2. Was the study described as double-blind? 3. Was there a description of withdrawals and drop-outs?	
Scoring the items ***Award a score of 1 point for each 'yes' and 0 points for each 'no'. There are no in-between marks.***	
Give 1 additional point if:	On question 1, the method of randomisation was described **and** it was **appropriate** (table of random numbers, computer-generated, coin tossing, etc.)
and/or if:	On question 2, the method of double-blinding was described **and** it was **appropriate** (identical placebo, active placebo, dummy, etc.)
Deduct 1 point if:	On question 1, the method of randomisation was described **and** it was **inappropriate** (patients were allocated alternatively, or according to date of birth, hospital number, etc.)
and/or if:	On question 2, the study was described as double-blind but the method of blinding was **inappropriate** (e.g. comparison of tablet vs. injection with no double-dummy)

TABLE 5 *Advice on using the scale*

1 Randomisation If the word randomised or any related words such as random, randomly, or randomisation are used in the report but the method of randomisation is not described, give a positive score to this item. A randomisation method is regarded as appropriate if it allows each patient the same chance of receiving each treatment but investigators could not predict which treatment was next. Methods of allocation using date of birth, date of admission, hospital numbers or alternation should not be regarded as appropriate.
2 Double-blinding A study must be regarded as double-blind if the term double-blind is used (even without describing the method) or if it is implied that neither care giver nor patient could identify the treatment being assessed.
3 Withdrawals and drop-outs Patients included in the study but who did not complete the observation period or who were not included in the analysis must be described. The number **and** the reasons for withdrawal must be stated. If there are no withdrawals, it should be stated. If there is no statement on withdrawals, a negative score (0 points) must be given.

aware that selective selection is still possible even when the group allocation is concealed until after consent has been obtained. The randomisation method is rated as inappropriate if the potential participants did not have the same chance of being included in any of the comparison groups (methods based on date of birth, hospital number or alternation). Even with excellent randomisation, selection bias may still be introduced if biased and selective withdrawals and drop-outs occur after the allocations have been made.[22] This is why an adequate description of withdrawals and drop-outs is included in the scale. With this information it is possible to analyse on an intention-to-treat basis (that is, all those randomised whether or not they were exposed to the study interventions[23]).

TABLE 6 *Scoring RCTs (maximum 5, minimum 1)*

	Score
Randomised?	
yes	1
Appropriate?	
yes (table)	1
no (alternative)	–1
Double-blind?	
yes	1
Appropriate?	
yes (double-dummy)	1
no	–1
Withdrawals described?	
yes	1

Rater bias can be minimised by blinding the person receiving the intervention, the individual administering it, the investigator measuring the outcome and the analyst. Blinding can be tested by asking the study patients and the researchers which intervention they had. This is not often done. The usual 'best' level of blinding is of the study subject and those making the observations (double-blinding). Double-blinding is often achieved by using control interventions with similar physical characteristics to those of the intervention under evaluation, or by the use of dummies when two or more interventions have to be given by different routes.

Sometimes, however, one of the interventions may produce effects which make blinding very difficult to sustain. Then the use of active placebos or active controls may decrease the likelihood of rater bias. All these precautions are relatively easy to achieve in pharmaceutical studies. In non-drug studies, testing under blind conditions is either difficult or inappropriate (e.g. surgical procedures) or impossible (e.g. acupuncture or TENS). The risk of rater bias limits the confidence with which conclusions can be reached. Studies which are not double-blind are known to risk an average exaggeration of treatment effect of 17% *(Box 1)*.[24]

BOX 1 Effect of randomisation and blinding on treatment effect[24]

	Overestimation	
	Inadequate	**Unclear**
Randomisation: concealment of treatment allocation	41%	30%
Blinding		
If not double-blind	17%	
Is this important in every setting? (unconscious)		
Is this practical in every setting? (surgery)		

References

1. Jadad AR, Moore RA, Carroll D, Jenkinson C, Reynolds DJM, Gavaghan DJ, *et al.* Assessing the quality of reports of randomized clinical trials: is blinding necessary? *Control Clin Trials* 1996;**17**:1–12.
2. Reiffenstein RJ, Schiltroth AJ, Todd DM. Current standards in reported drug trials. *Can Med Assoc J* 1968;**99**:1134–5.
3. Chalmers TC, Smith H, Blackburn B, Silverman B, Schroeder B, Reitmand D, *et al.* A method for assessing the quality of a randomized control trial. *Control Clin Trials* 1981;**2**:31–49.
4. Tyson JE, Furzan JA, Reisch IS, Mize SG. An evaluation of the quality of therapeutic studies in perinatal medicine. *J Pediatr* 1983;**102**:10–13.
5. Andrew A. Method for assessing of the reporting standard of clinical trials with roentgen contrast media. *Acta Radiol Diag* 1984;**25**:55–8.
6. Evans M, Pollock AV. A score system for evaluating random control clinical trials of prophylaxis of abdominal surgical wound infection. *Br J Surg* 1985;**72**:256–60.
7. Poynard T. Evaluation de la qualité méthodologique des essais thérapeutiques randomisés. *Presse Med* 1988;**17**:315–18.
8. Chalmers I, Hetherington J, Elbourne D, Keirse MJNC, Enkin M. Materials and methods used in synthesizing evidence to evaluate the effects of care during pregnancy and childbirth. In: Chalmers I, Enkin M, Keirse MJNC, editors. Effective care in pregnancy and childbirth. Oxford: Oxford University Press 1989;38–65.
9. Gøtzsche PC. Methodology and overt and hidden bias in reports of 196 double blind trials of non-steroidal anti-inflammatory drugs in rheumatoid arthritis [published erratum appears in *Control Clin Trials* 1989;**50**:356]. *Control Clin Trials* 1989;**10**:31–56.
10. Reisch JS, Tyson JE, Mize SG. Aid to the evaluation of therapeutic studies. *Pediatrics* 1989;**84**:815–27.
11. ter Riet G, Kleijnen J, Knipschild P. Acupuncture and chronic pain: a criteria-based meta-analysis. *J Clin Epidemiol* 1990;**43**:1191–9.
12. Brown SA. Measurement of quality of primary studies for meta-analysis. *Nurs Res* 1991;**40**:352–5.
13. Koes BW, Bouter LM, Beckerman H, van der Heijden, Knipschild PG. Physiotherapy exercises and back pain: a blinded review [see comments]. *BMJ* 1991;**302**:1572–6.
14. Koes BW, Assendelft WJ, van der Heijden, Bouter LM, Knipschild PG. Spinal manipulation and mobilisation for back and neck pain: a blinded review [see comments]. *BMJ* 1991;**303**:1298–303.
15. Detsky AS, Naylor CD, O'Rourke K, McGeer AJ, l'Abbé KA. Incorporating variations in the quality of individual randomized trials into meta-analysis. *J Clin Epidemiol* 1992;**45**:255–65.
16. Nurmohamed MT, Rosendaal FR, Buller HR, Dekker E, Hommes DW, Vandenbroucke JP, *et al.* Low molecular-weight heparin versus standard heparin in general and orthopaedic surgery: a meta-analysis. *Lancet* 1992;**340**:152–6.
17. Oxman AD, Guyatt GH. Validation of an index of the quality of review articles. *J Clin Epidemiol* 1991;**44**:1271–8.
18. Fink A, Kosecoff J, Chassin M, Brook RH. Consensus methods: characteristics and guidelines for use. *Am J Public Health* 1984;**74**:979–83.

19. Feinstein AR. Clinimetrics. Newhaven: Yale University Press, 1987.

20. Fleiss JL, Gross AJ. Meta-analysis in epidemiology, with special reference to studies of the association between exposure to environmental tobacco smoke and lung cancer: a critique. *J Clin Epidemiol* 1991;**44**:127–39.

21. Chalmers I. Evaluating the effects of care during pregnancy and childbirth. In: Chalmers I, Enkin M, Keirse MJNC, editors. Effective care in pregnancy and childbirth. Oxford: Oxford University Press, 1989;1–37.

22. Sackett DL, Gent M. Controversy in counting and attributing events in clinical trials. *N Engl J Med* 1979;**301**:1410–12.

23. Peto R, Pike MC, Armitage P, Breslow NE, Cox DR, Howard SV, *et al.* Design and analysis of randomised clinical trials requiring prolonged observation of each patient. 1. Introduction and design. *Br J Cancer* 1976;**34**:585–612.

24. Schulz KF, Chalmers I, Hayes RJ, Altman DG. Empirical evidence of bias: dimensions of methodological quality associated with estimates of treatment effects in controlled trials. *JAMA* 1995;**273**:408–12.

Chapter 4

Pain measurement, study design and validity

The efficacy of analgesic interventions is judged by the change they bring about in the patient's report of pain. A brief description of methods of pain measurement follows. In the second part of this chapter, the problems and some solutions when using pain measurement data for systematic reviews are discussed.

Pain measurement

Pain is a personal experience which makes it difficult to define and measure. It includes both the sensory input and any modulation by physiological, psychological and environmental factors. Not surprisingly, there are no objective measures – there is no way to measure pain directly by sampling blood or urine, or by performing neurophysiological tests. Measurement of pain must therefore rely on recording the patient's report. The assumption is often made that because the measurement is subjective it must be of little value. The reality is that, if the measurements are made properly, remarkably sensitive and consistent results can be obtained. There are contexts, however, in which it is not possible to measure pain at all, or when reports are likely to be unreliable. These include impaired consciousness, young children, psychiatric pathology, severe anxiety, unwillingness to cooperate, and inability to understand the measurements. Such problems are deliberately avoided in trials.

Measurement scales

Most analgesic studies include measurements of pain intensity and/or pain relief, and the commonest tools used are categorical and visual analogue scales (VAS).

Categorical and visual analogue scales

Categorical scales use words to describe the magnitude of the pain and were the earliest pain measure.[1] The patient picks the most appropriate word that describes a pain. Most research groups use four words (none, mild, moderate and severe). Scales to measure pain relief were developed later. The most common is the five category scale (none, slight, moderate, good or lots, and complete).

For analysis purposes, numbers are given to the verbal categories:

- **pain intensity** – none = 0, mild = 1, moderate = 2, severe = 3
- **pain relief** – none = 0, slight = 1, moderate = 2, good or lots = 3, complete = 4.

Data from different subjects is then combined to produce means (rarely medians) and measures of dispersion (usually standard errors of means). The validity of converting categories into numerical scores was checked by comparison with concurrent VAS measurements. Good correlation was found, especially between pain relief scales using cross-modality matching techniques.[2–4] Results are usually reported as continuous data, mean or median pain relief or intensity. Few studies present results as discrete data, giving the number of participants who report a certain level of pain intensity or relief at any given assessment point. The main advantages of the categorical scales are that they are quick and simple. However, the limited number of descriptors may force the scorer to choose a particular category when none of them describes the pain satisfactorily.

The VAS, a line with left-hand end labelled 'no relief of pain' and right-hand end labelled 'complete relief of pain' *(Figure 6),* seem to overcome this limitation. Patients mark the line at the point which corresponds to their pain. The scores are obtained by measuring the distance, usually in millimetres, between the no relief end and the patient's mark. The main advantages of VAS are that they are simple and quick to score, avoid imprecise descriptive terms and provide many points from which to choose. More concentration and coordination are needed, however, which can be difficult postoperatively or with neurological disorders.

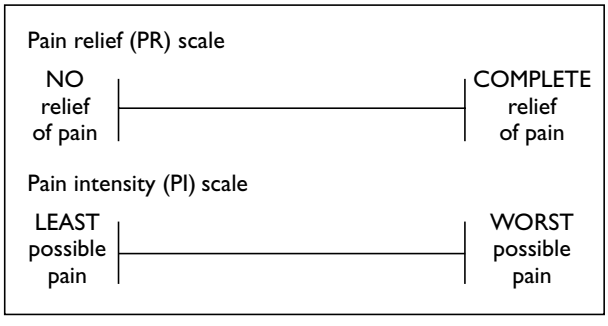


FIGURE 6 *Visual analogue scales*

Pain relief scales are perceived as more convenient than pain intensity scales, probably because patients have the same baseline relief (zero) whereas they could start with different baseline intensity (usually moderate or severe). Relief scale results are thus easier to compare. They may also be more sensitive than intensity scales.[4,5] A theoretical drawback of pain relief scales is that the patient has to remember what the pain was like to begin with.

Other tools

Verbal numerical scales and global subjective efficacy ratings are also used. Verbal numerical scales are regarded as an alternative or complementary to the categorical and VAS scales. Patients choose a number for the pain intensity or relief (for pain intensity, 0 usually represents no pain and 10 the maximum possible, and for pain relief, 0 represents none and 10 complete relief). They are very easy and quick to use, and correlate well with conventional VASs.[6]

Global subjective efficacy ratings, or simply global scales, are designed to measure overall treatment performance. Patients are asked questions such as, 'How effective do you think the treatment was?' and answer using a labelled numerical or a categorical scale. Although these judgements probably include adverse effects they can be the most sensitive in discriminating between treatments. One of the oldest scales was the binary question, 'Is your pain half gone?'. It has the advantage that it has a clearer clinical meaning than a 10 mm shift on a VAS. The disadvantage, for the small trial intensive measure pundits at least, is that all the potential intermediate information (1–49% or > 50%) is discarded.

Analgesic requirements (including patient-controlled analgesia, PCA), special paediatric scales, and questionnaires (such as McGill) are also used. The limitation to guard against is that they usually reflect other experiences as well as or instead of pain.[7]

Judgement by the patient rather than by the carer is the ideal. Carers tend to overestimate the pain relief compared with the patient.

Analysis of scale results – summary measures

In the research context, pain is usually assessed before the intervention is made and then on multiple occasions. Ideally, the area under the time–analgesic effect curve for the intensity (sum of pain intensity differences, SPID) or relief (total pain relief; TOTPAR) measures is derived.

$$\text{SPID} = \sum_{t=0-6}^{n} \text{PID}_t \qquad \text{TOTPAR} = \sum_{t=0-6}^{n} \text{PR}_t$$

Where at the tth assessment point ($t = 0, 1, 2,..., n$), P_t and PR_t are pain intensity and pain relief measured at that point, respectively, P_0 is pain intensity at $t = 0$ and PID_t is the pain intensity difference calculated as $(P_0 - P_t)$.

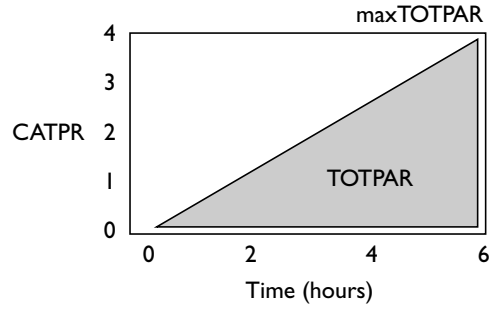
These summary measures reflect the cumulative response to the intervention. Their disadvantage is that they provide no information about the onset and peak of the analgesic effect. If onset or peak are important, then time to maximum pain relief (or reduction in pain intensity) or time for pain to return to baseline are necessary.

Using pain measurement data for systematic reviews

Standardising the summary measures

The method used to standardise TOTPAR values, derived from a categorical verbal rating scale of relief (CATPR), is shown in *Figure 7*. The actual TOTPAR value is divided by the maximum possible TOTPAR score (maximum duration in hours multiplied by the maximum pain relief score) and converted to a percentage.

This calculation presumes that categorical relief score data is available. One major problem we faced is that not all trials use this classic scale. In



Categorical verbal rating scale pain relief (CATPR):
0 = none, 1 = slight, 2 = moderate, 3 = good, 4 = complete

$$\frac{\text{TOTPAR}}{\text{maxTOTPAR}} \times 100 = \%\ \text{maxTOTPAR}$$

FIGURE 7 *Calculating percentage of maximum possible pain relief score*

order to include trials which used different scales, ways of converting those different scales back to the common denominator of % maxTOTPAR had to be developed. The development and validation of these methods is discussed in chapter 5. We still cannot include trials which use analgesic drug consumption (e.g. PCA), or trials which use non-standard scales.

The hazard for meta-analysis is that if many papers have to be discarded because they do not use standard scales, are the remaining trials representative? For most drug interventions this has not proved to be a major problem, because the majority of trials used standard methods. However, there are exceptions. For some older drugs, such as dihydrocodeine, remarkably few trials were found which used standard methods. For academically inspired investigations, as opposed to trials required for drug registration, many trials which use non-standard methods have had to be excluded.

Restricting to moderate and severe initial pain intensity

The trail blazers of analgesic trial methodology found that if patients had no pain to begin with, it was impossible to assess analgesic efficacy, because there was no pain to relieve. To optimise trial sensitivity, a rule was developed – only those patients with moderate or severe pain intensity at baseline would be studied. Patients with mild pain or no pain would not.

This study has stayed true to this rule; trials of a given intervention have been excluded if the trials studied patients with mild or no initial pain. As with exclusions because of non-standard methods, there have been few pharmacological trials where the rule on baseline pain has led to exclusion but for pre-emptive techniques and local anaesthetic blocks it has been a major problem.

How do you know what is moderate or severe pain on a pain intensity VAS?

The usual criterion to ensure adequate sensitivity for analgesic trials is to test the intervention on patients who have established pain of moderate to severe intensity. When a VAS is the only pain measure in a trial we need to know what point on it represents moderate pain, so that the trial can be included in a meta-analysis with an inclusion criterion of baseline pain of at least moderate intensity.

In order to answer this question, individual patient data from 1080 patients from RCTs of various analgesics were used.[8] Baseline pain had been measured using both a four-point categorical pain intensity scale and a VAS pain intensity scale. The distribution of the VAS scores was checked for 736 patients reporting moderate pain and for 344 reporting severe pain. The VAS scores corresponding to moderate or severe pain were also checked by gender.

Baseline VAS scores for patients reporting moderate pain were significantly different from those of patients reporting severe pain (see *Table 7* and *Figure 8)*. Of the patients reporting moderate pain, 85% scored over 30 mm on the corresponding VAS, with a mean score of 49 mm. For those reporting severe pain, 85% scored over 54 mm with a mean score of 75 mm. There was no difference between the corresponding VAS scores for men and women. These results indicate that if a patient records a baseline VAS score in excess of 30 mm they would probably have recorded at least moderate pain on a four-point categorical scale.

TABLE 7 *Descriptive statistics for the distribution of VAS pain intensity scores*

	Baseline Pain (using a 4-point categorical scale)	
	Moderate	**Severe**
n	736	344
Mean (mm)	49	75
Standard deviation	17	18
Median (mm)	49	76
90% patients > (mm)	26	49
85% patients > (mm)	30	54

Study design and validity

Pain measurement is one of the oldest and most studied of the subjective measures, and pain scales have been used for over 40 years. Even in the early days of pain measurement there was understanding that the design of studies contributed directly to the validity of the result obtained. Trial designs which lack validity produce information that is, at best, difficult to use and, at worst, useless.

Placebo

People in pain respond to placebo treatment. Some patients given placebo obtain 100% pain relief (see chapter 5). The effect is reproducible, and some work has been done to try and assess the characteristics of the 'placebo responder' by gender, race

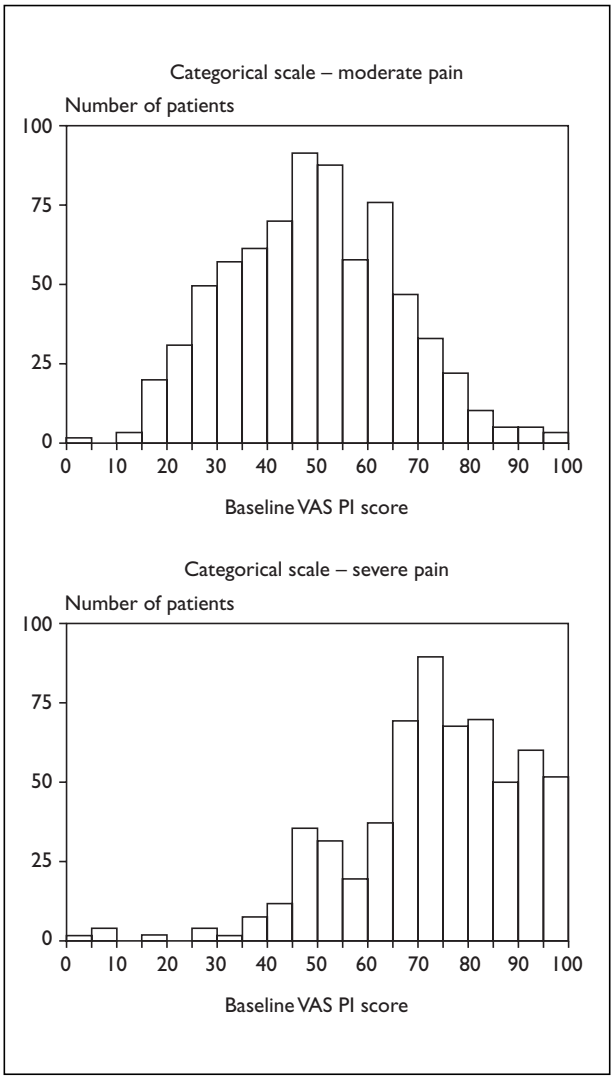


FIGURE 8 *Frequency distributions of initial VAS pain intensity scores by initial categorical verbal pain intensity*

and psychological profile. None has succeeded but women are known to respond better than men to some analgesics, getting more analgesia from the same plasma concentration of drug.

RCT

Because the placebo response was an established fact in analgesic studies, randomisation was used early in studies to try to avoid any possibility of bias from placebo responders, and to equalise their numbers in each treatment group. This was true even in studies without placebo, since an excess of placebo responders in an active treatment arm of a study might inflate the effects of an analgesic.

Sensitivity

Particularly for a new analgesic, a trial should prove its internal sensitivity – that is that the study is an adequate analgesic assay. This can be done in several ways. For instance, if a known analgesic (such as paracetamol) can be shown to have statistical difference from placebo, then the analgesic assay should be able to distinguish another analgesic of similar effectiveness. Alternatively, two different doses of a standard analgesic (such as morphine) could be used – showing the higher dose to be statistically superior to the lower dose again provides confidence in the sensitivity of the assay.

Failure to demonstrate sensitivity in one assay invalidates the results from that particular assay. However, the results could still be included in meta-analysis.

Equivalence

Studies of analgesics of an A versus B design are notoriously difficult to interpret *(Figure 9)*. If there is a statistical difference, then that suggests sensitivity. Lack of a significant difference *(Figure 9,* top graph) means nothing – there is no method by which to determine whether there is an analgesic effect which is no different between A and B, or whether the assay lacks the sensitivity to measure a difference that is actually present.

This is not just a problem for pain studies.[9,10] Designs which minimise these problems include using two doses of a standard analgesic plus placebo to establish sensitivity *(Figure 9,* middle and bottom graphs). Simple calculations could show what dose of the new analgesic was equivalent to the usual dose of the standard analgesic.

Problems

The correct design of an analgesic trial is situation-dependent. In some circumstances, very complex designs have to be used to ensure sensitivity and validity.

No gold standard

There may be circumstances in which there is no established analgesic treatment of sufficient effectiveness to act as a gold standard against which to measure a new treatment, as is often the case in chronic pain. Clearly, the use of placebo or no-treatment controls is of great importance, especially when effects are to be examined over prolonged periods of weeks or months.

However, paradoxically, it is these very circumstances in which ethical constraints act against using placebo or non-treatment controls because of the need to do **something**. In acute pain studies, conversely, there are few problems with using placebos, since the failure of a placebo (or any treatment) can be dealt with by prescribing additional analgesics which should work.

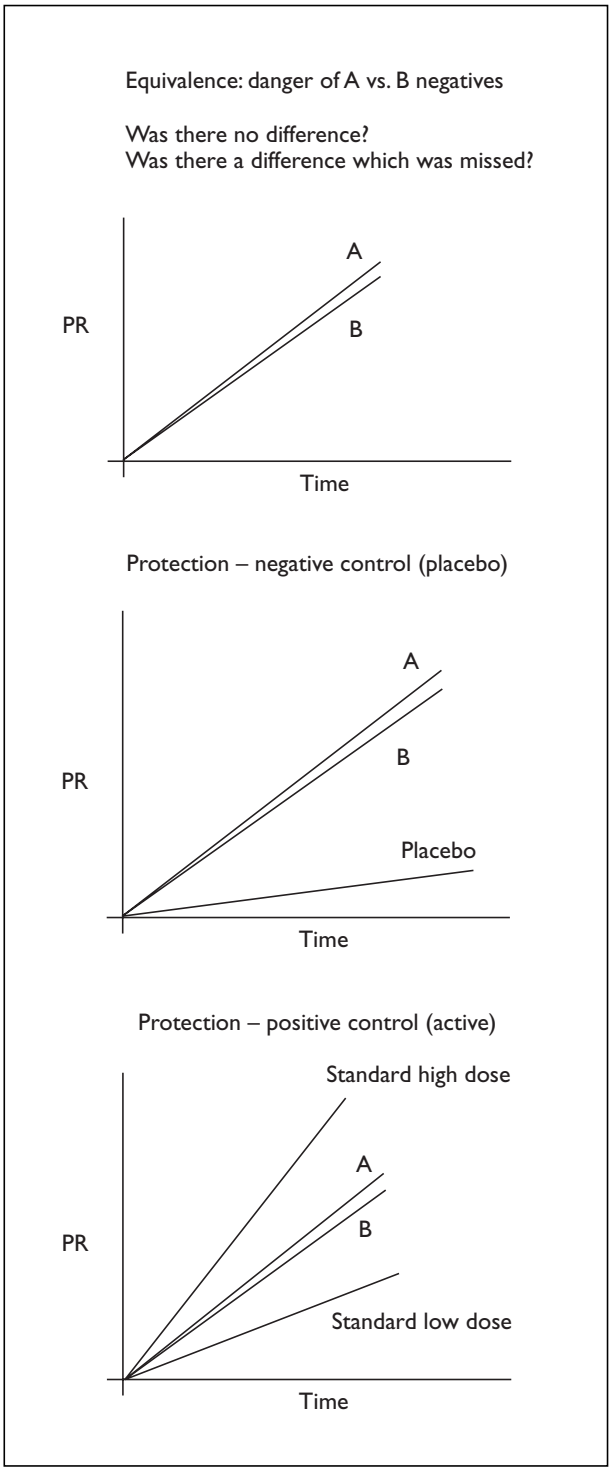


FIGURE 9 *Using placebo or active comparators to protect against A vs. B negative results*

When there is no pain to begin with

Clearly, where there is no pain it is difficult to measure an analgesic response. Yet a number of studies seek to do this by pre-empting pain, or by using an intervention where there is no pain (intraoperatively, for instance) to produce analgesia when pain is to be expected.

These are difficult but not impossible circumstances in which to conduct research. Meticulous attention to trial design is necessary to be able to demonstrate differences.

References

1. Keele KD. The pain chart. *Lancet* 1948;**2**:6–8.
2. Scott J, Huskisson EC. Graphic representation of pain. *Pain* 1976;**2**:175–84.
3. Wallenstein SL, Heidrich IG, Kaiko R, Houde RW. Clinical evaluation of mild analgesics: the measurement of clinical pain. *Br J Clin Pharmacol* 1980;**10**:319–27S.
4. Littman GS, Walker B.R., Schneider BE. Reassessment of verbal and visual analogue ratings in analgesic studies. *Clin Pharmacol Therap* 1985;**38**:16–23.
5. Sriwatanakul K, Kelvie W, Lasagna L. The quantification of pain: an analysis of words used to describe pain and analgesia in clinical trials. *Clin Pharmacol Ther* 1982;**32**:141–8.
6. Murphy DF, McDonald A, Power C, Unwin A, MacSullivan R. Measurement of pain: a comparison of the visual analogue with a nonvisual analogue scale. *Clin J Pain* 1988;**3**:197–9.
7. Jadad AR, McQuay HJ. The measurement of pain. In: Pynsent P, Fairbank J, Carr A, editors. Outcome measures in orthopaedics. Oxford: Butterworth Heinemann, 1993:16–29.
8. Collins SL, Moore RA, McQuay HJ. The visual analogue pain intensity scale: what is moderate pain in millimetres? *Pain* 1997;**72**:95–7.
9. Jones B, Jarvis P, Lewis JA, Ebbutt AF. Trials to assess equivalence: the importance of rigorous methods. *BMJ* 1996;**313**:36–9.
10. McQuay H, Moore A. Placebos are essential when extent and variability of placebo response are unknown. *BMJ* 1996;**313**:1008.

Chapter 5

Estimating relative effectiveness

Minor analgesics, such as paracetamol, ibuprofen and combinations with opiates like codeine or dextropropoxyphene, are used often to treat pain. There are few direct comparisons, one with another, but most trials contain a placebo, which has the potential to be the universal comparator. Instead of measuring relative effectiveness through multiple comparisons of different drugs, it should be possible to compare the absolute effectiveness of analgesics against placebo.

Some aspects of clinical trial methods relating to the placebo response in clinical trials of single doses of analgesics using classical methods are examined in this chapter. The ways in which data can be extracted from published studies for use in meta-analysis are then determined.

Placebo responses in analgesic trials

The placebo response is confusing.[1] Two common misconceptions are that a fixed fraction (one-third) of the population responds to placebo and that the extent of the placebo reaction is also a fixed fraction (again about one-third of the maximum possible[2]). As Wall points out, these ideas stem from a misreading of Beecher's work of 40 years ago.[1]

In Beecher's five acute pain studies, 139 patients (31%) of 452 given placebo had 50% or more relief of postoperative pain at two checked intervals.[3] The proportion of patients who had 50% or more relief of pain varied across the studies, ranging from 15% to 53%. There was neither a fixed fraction of responders, nor a fixed extent of response.

Placebo responses have also been reported as varying systematically with the efficacy of the active analgesic medicine. Evans pointed out that in seven studies the placebo response was always about 55% of the active treatment, whether that was aspirin or morphine: the stronger the drug, the stronger the placebo response.[4]

Randomised, double-blind trials are meant to eliminate (or at least minimise) both selection and observer bias; Evans' observation suggests that significant observer bias occurs. Wall[1] rightly questions the blindness of these trials if this result was correct, and elegantly dissects the areas where 'leakage' of blinding can occur (patient–patient, patient–doctor, patient–nurse).

Both these observations call into question the validity of the methods used to gather the data. If the methods are faulty, how reliable are the answers? Therefore, in the first part of this chapter the nature of the variation in placebo responses in five randomised, double-blind, parallel-group trials in postoperative pain is examined, together with the relationship of the variation to the analgesic effectiveness of the active treatments.

Methods

Individual patient data was used from five placebo-controlled double-blind RCTs, performed over a 10-year period by the Pain Research Group in Oxford, in which the analgesic effects of various drugs in postoperative pain were investigated.[5–9] All were randomised, double-blind and parallel-group trials of single doses of drugs given orally. Randomisation was made using random number tables. Drugs were prepared outside the hospital in which the studies were done. Treatment codes were not broken until the studies were finished. All drugs used within a study were identical. Drugs were given in a standardised way by the nurse observer. The methods used by the trained nurse observers to measure pain were identical. Patients were asked a standardised battery of questions in a fixed order at each assessment point in the studies. All patients knew that a placebo was one of several possible treatments. All patients had moderate or severe pain within 72 hours of their operations, and all were aware that they could withdraw from the study at any time for any reason.

Each study used five scales for pain; three for pain intensity and two for pain relief. Of these the five-point CATPR scale for pain relief (0 = none, 1 = slight, 2 = moderate, 3 = good, 4 = complete) was chosen for this analysis because it was closest to Beecher's original method. For each patient the area under the curve of pain relief (categorical scale) against time was calculated (TOTPAR). The percentage of the maximum possible for this summary measure was then calculated (% maxTOTPAR).[10]

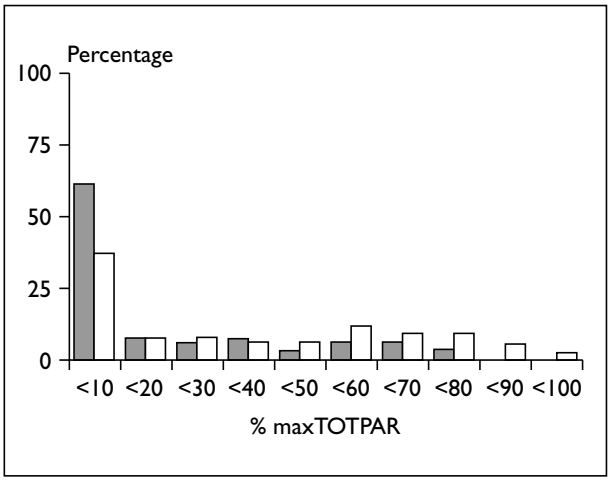


FIGURE 10 *Distribution of % maxTOTPAR scores for the 130 patients given placebo (■) and for the 395 patients given active drugs (□)*

Results

In the five trials 130 patients had a placebo. Individual patients' scores with placebo varied from 0 to 100% of the maximum possible pain relief.

The distribution of these % maxTOTPAR scores is shown in *Figure 10*. In the five trials, 395 patients were given active drugs. Individual patient scores with different active drugs varied from 0% to 97% of the maximum possible pain relief. The distribution of these % maxTOTPAR scores for the active drugs is shown in *Figure 10*.

The mean % maxTOTPAR scores for the five placebo groups varied from 11% to 29%, and the mean scores for the active drugs varied from 12% to 49%. The relationship between the mean scores for the active drugs and the mean placebo scores is shown in *Figure 11*. Mean placebo scores were related to the mean score for the active drugs in each trial such that the higher the active score, the higher the placebo score. A similar relationship obtained for the best and for the worst active drug from each of the five trials. On average, the mean placebo results were 54% of the mean active results based on a slope of 0.54; 95% confidence intervals (CIs) around the slope, 0.03–1.08.

The relationship between the median scores for active and placebo treatments is also shown in *Figure 11*. There was little relationship between the two and, on average, the median placebo score was less than 10% of the median active drug score. The slope to the regression line was 0.12 (95% CIs, –0.24, 0.48) and included no relationship between placebo response and extent of the response to active analgesic.

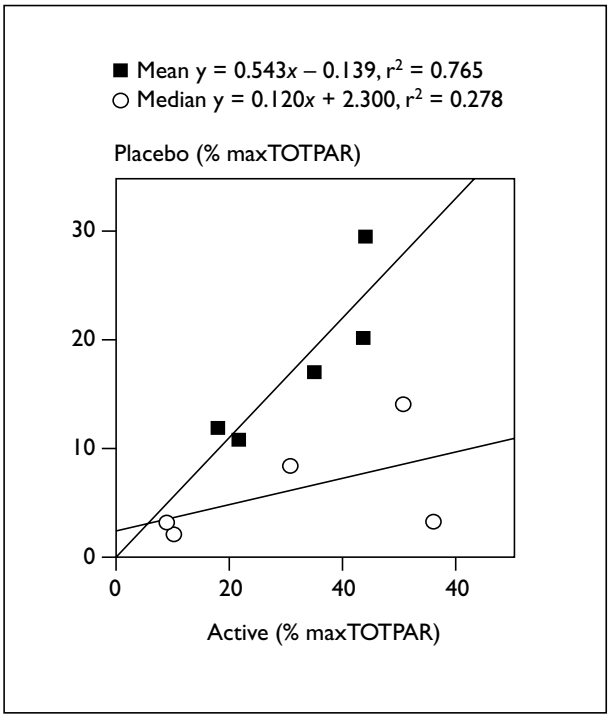


FIGURE 11 *% maxTOTPAR placebo and active scores for five trials*

The same pattern of results was also found when the analysis was repeated using the results from the VAS for pain relief.

Comment

The variation of the placebo response in the acute pain setting found by Beecher some 40 years ago is confirmed by these results. Using the dichotomous measure of greater than 50% pain relief at 45 and 90 minutes, Beecher found that a range of 15–53% of patients given placebo had better than 50% relief in five acute pain studies.[3] Here, using the derived dichotomous measure of 50% maximum pain relief, a range of 7–37% of patients given placebo achieved better than 50% relief across the five studies (see *Table 8*).

In analgesic trials the response of a group of patients to a treatment is usually described not as a dichotomous variable (like the proportion of patients with at least 50% relief), but rather as a continuous variable (the mean extent of the response). The common description of pain intensity difference or pain relief is thus as the mean with standard deviations or standard errors of the mean, as if the data were normally distributed.

Patient responses were not normally distributed, either for patients given placebo or for those given an active treatment *(Figure 10)*. The predominant

TABLE 8 *Results with placebo; mean (SD), median (interquartile range) and number of patients with > 50% of % maxTOTPAR in the five studies*

Study	No.	Mean % maxTOTPAR (SD)	Median % maxTOTPAR (interquartile range)	Number of patients with > 50% of % maxTOTPAR	% of patients with at least 50% of % maxTOTPAR
Porter, *et al*.[5]	21	11.9 (19.3)	3.1 (16.4)	2	10
Evans, *et al*.[6]	30	29.4 (29.1)	14.0 (53.0)	11	37
McQuay, *et al*.[7]	19	20.1 (29.1)	3.1 (27.3)	4	21
McQuay, *et al*.[8]	30	10.7 (17.8)	2.1 (8.3)	2	7
McQuay, *et al*.[9]	30	16.9 (21.2)	8.3 (25.0)	2	7

group was that getting less than 10% of maximum relief – 62% of patients given placebo and 37% of those given an active treatment. In these circumstances, the use of a mean as a description is not valid, and the use of a median is more sensible. Averaging results to describe them is a historical hangover.

In describing the placebo groups, therefore, the range of mean placebo response of 11–29% of maximum *(Table 8)* becomes a range of median placebo response of 2–14% and a range of the proportion of patients with at least 50% of % maxTOTPAR of 7–37%. Regressing median placebo response against median active response from the same five trials yielded a poor correlation, with a regression line no different from the horizontal, which would be the expected result if there was no bias. The idea that there is a constant relationship between active analgesic and placebo response is therefore an artefact of using an inappropriate statistical description.

It is the comparison of the mean data from placebo and active treatments that led to the observation[4] that placebo is about 55% as effective as an active treatment, whatever active treatment is used. In the five trials here, comparison of the mean placebo response with the mean active treatment *(Figure 11)* produced a regression with a slope of 0.54 – exactly the same result!

This defies logic unless there was considerable bias despite randomisation and the use of double-blind methods, and would, if true, undermine the confidence placed in analgesic trial results. But is it true?

Randomisation controls for selection bias, and the double-blind design is there to control observer bias. Patients knew a placebo was one possible treatment, and the investigators knew the study design and active treatments; it has been suggested that this can modify patients' behaviour.[11,12] A small number of patients may have had opportunities to communicate with each other. Doctors who knew the trial design obtained consents from patients, and this may also be a source of bias.[13] The nurse observer spent most time with the patients, but in standardised situations. This would be the most likely source of bias, as the nurse might be able to influence a patient's response by his/her demeanour based on experience of other patients' reactions. That would produce time-dependent changes in study results that have been observed before.[14]

Bias may still occur but its effects are slight, and this has important consequences. It means that results obtained over a range of clinical conditions and times may be combined in meta-analyses with confidence. Gøtzsche has confirmed similar magnitudes of effect for NSAIDs in active and placebo-controlled studies,[15] showing that the presence of a placebo does not affect the active treatment – the alternative hypothesis.

Deriving dichotomous outcome measures from continuous data in RCTs of analgesics

The problem is that, in most published trial reports, the only value available which describes the magnitude of analgesic effect is the mean and standard deviation of the SPID or TOTPAR. Is it possible, then, to use this to generate other, more useful data with which meta-analysis can work with confidence? Meta-analytic outcomes using mean values from different trials have been explored,[16,17] but the result is a complicated analysis which is not intuitively accessible to doctor or patient. If

individual patient information was available from every RCT of analgesics, dichotomous data could be extracted for NNT calculations. The reality is that individual patient data is not available, so that the problem is how to derive dichotomous outcomes from the published mean data. A full version of these arguments is published elsewhere.[18]

A proposed solution

The hypothesis was examined that, in pharmacological interventions in acute pain:

(i) a relationship exists between the descriptive mean value for pain relief and a dichotomous description of the same data set
(ii) knowing this relationship allows the conversion of descriptive mean values for pain relief into dichotomous data that can be used with confidence for meta-analysis.

Relationships that exist between treatment group means and some simple extractable variables from a known data set are an obvious place to start. What is required as an extractable variable is a single value, for instance, the proportion or number of patients who have achieved 50% pain relief. If treatment group means reliably predict the proportion with 50% pain relief, this suggests that the relationship between the two variables is a product of the underlying distribution. One benefit of using the proportion of patients who have achieved at least 50% pain relief is that it is clinically intuitive.

The robustness of such a relationship can be tested in various ways. The gold standard would be to test relationships between mean and dichotomous variables developed from one set of trials using data from other trials; however, this has not proved possible.

In the absence of available information from real trials, surrogate trials can be obtained through simulation. Simulation methods have been used to generate individual patient data for large numbers of trials using the underlying distribution from randomised trials of pharmacological interventions performed in Oxford over about 15 years using standard methods. This approach generates precision in defining the underlying distribution of the data, and tests the assumptions made in deriving the technique for converting mean pain relief data into dichotomous data.

While simulation methods can give a degree of confidence that the general approach has validity, it is testing against other, real, data sets that will allow the method to be used in meta-analysis.

Methods for converting mean to dichotomous data from clinical trials of analgesics, given in single doses using classical analgesic methodology, have been determined in three stages, all of which use at least 50% maxTOTPAR as a final dichotomous outcome.

Stage 1 Use of Oxford data from about 1500 patients combined with mathematical modelling using TOTPAR scales.
Stage 2 Verification with an external data set of 3500 patients using TOTPAR scales.
Stage 3 Examination of the use of other scales.

For each stage, methods and results are shown separately and then discussed together.

Stage 1 methods

Actual patient data

Individual patient data were taken from 12 placebo- and active-controlled double-blind randomised trials in which the analgesic effects of various drugs in postoperative pain were investigated.[5–9,19–26] The trials were undertaken over a 15-year period by the Pain Research Group, Oxford. Complete individual patient information over 4 or 6 hours was available for a number of pain and pain relief scales. All drugs were administered orally, except sublingual buprenorphine[20] and intramuscular opioids.[26]

All the studies were randomised, double-blind and parallel-group. Patients were told about the study by the nurse observer the day before surgery. Informed consent was obtained by the doctor that evening. Random number tables were used for randomisation. Drugs were prepared outside the hospital in which the studies were undertaken. Treatment codes were not broken until the studies were completed. All drugs within a study were identical in appearance and double-dummy methods were used when different routes of administration were compared. Drugs were given in a standardised way by the nurse observer. The methods used by the trained nurse observers to measure pain were identical.

Patients were asked a standardised battery of questions in a fixed order at each assessment time in the studies. In placebo-controlled trials, all patients knew that placebo was one possible treatment. All patients had moderate or severe pain within 72 hours postoperatively, and all were aware that they could withdraw from the study at any time for any reason. At the start of the assessments the nurse observer made sure that patients had recovered sufficiently from the anaesthetic and were able to communicate reliably. Studies with more than one nurse observer were block randomised, with

one nurse responsible for each block. Only one nurse assessed any one patient. If no pain relief was obtained from the test medication by 1 hour, or if the pain intensity subsequently reverted to the initial value before the end of the 6-hour study, patients were given analgesia ('escape analgesia').

Each study used five scales for pain; three for pain intensity and two for pain relief. In this study, the categorical measurement of pain relief with a five-point categorical verbal rating scale (CATPR: 0 = none, 1 = slight, 2 = moderate, 3 = good, 4 = complete) was used because it has been shown that with this scale placebo responses are independent of active treatment efficacy.[27]

For each patient, the area under the curve of pain relief (categorical scale) against time was calculated (TOTPAR). The percentage of the maximum possible for this summary measure was then calculated (% maxTOTPAR), as well as the numbers and proportion of each group with at least 50% maxTOTPAR (or percentage > 50% maxTOTPAR to accommodate unequal group sizes). The dichotomous descriptor of at least 50% maxTOTPAR was chosen because it is a simple clinical endpoint of pain half relieved, easily understood by professionals and patients.

The relationship between the mean % maxTOTPAR and the actual number of patients with at least 50% maxTOTPAR was examined by linear regression analysis. Using the equation to the regression line, the calculated number of patients with at least 50% maxTOTPAR was then compared with the actual number.

Odds ratios and their 95% CIs were calculated from standard formulae incorporating a fixed-effects model and NNTs, using the method of Cook and Sackett.[28] Where the same treatment (placebo or active) had been given in different trials, data from individual treatment arms were combined.

Simulations

The underlying distribution using % maxTOTPAR for individual real patients in the actual 45 treatments was approximately uniform over the range 10–100% of % maxTOTPAR, with a spike in the range 0–10% of % maxTOTPAR. This was an amalgamation of patient data from all the treatments and was unlikely to reflect the actual distribution within any one treatment.

Because the possibility exists that statistical differences in distribution could occur in treatment arms with relatively small patient numbers, simulations were conducted to test how robust the relationships developed with actual treatments and real patients might be. The simulations had three main aims, as follows.

1. To generate a very large number of simulated active treatments (10,000) with a mean of 30 simulated patients (standard deviation, 3 patients, minimum group size 15 patients) in each, where the % maxTOTPAR for each simulated patient was generated randomly from a distribution similar to the real data. Comparable results from real and simulated data would allow the conclusion that the conversion technique was dependent only on the amalgamated distribution of % maxTOTPAR from all trials, and not on the underlying distribution of % maxTOTPAR within each trial.
2. To show that, for each simulated treatment, mean % maxTOTPAR could be converted to the calculated number with at least 50% maxTOTPAR, using the techniques developed for the 45 actual treatments, and, for these simulated treatments, to compare the calculated number with at least 50% maxTOTPAR with the number generated in the simulation. This would provide an indication of how accurate the conversion technique was likely to be for a large data set with this underlying distribution.
3. To generate simulated individual patient data using two different underlying distributions (normal distribution and a uniform distribution, ensuring in each case that the mean was similar to that for the real data), in order to test the extent to which the accuracy of the conversion technique was dependent on the underlying distribution.

Computer codes were written in Fortran and run on the Oxford University DEC Vax Cluster. Uniform random numbers in the range 0–1 ($U[0,1]$) were obtained using the intrinsic function 'ran', and these were then used to calculate both random treatment sizes and individual patient data with the appropriate underlying distribution, as described below.

(i) Treatment sizes were assumed to be normally distributed with a mean of 30 and a standard deviation of 3. These were calculated by transforming the $U[0, 1]$ values into normal values with the required mean and standard deviation using the Box–Mueller algorithm.[29] If any generated value of the group size was below

15 it was discarded and a new value generated which fell within the appropriate range.

(ii) For generation of the 'simulated actual' distribution, the U [0, 1] value generated was first multiplied by 140 (giving a U [0, 140] distribution), but for any values greater than 100 the value was discarded and a new value generated which was multiplied by 10. This process ensured that 50/140 (36%) patients were uniformly distributed in the range 0–10% maxTOTPAR while the remaining 64% were uniformly distributed in the range 10 to 100. Standard techniques were then used to show that a distribution generated in this way had a theoretical mean of 37.1 and a standard deviation of 31.7.

(iii) For generation of the 'normal' distribution, the Box–Mueller algorithm was again used to generate the appropriate values but, in this case, it was necessary to restrict generated values to the range 0–100% maxTOTPAR. Since this restriction process altered the mean of the underlying distribution, the appropriate values to be used in the simulation to give a mean of 37 were determined by iteration.

(iv) For the generation of the 'uniform' distribution, the value U [0, 1] was multiplied by 74.0 to obtain a distribution which was uniform on [0, 74], with a mean of 37.

Stage 1 results

The actual trials used in the analysis, the treatments used, numbers in each group, mean % maxTOTPAR and the numbers of patients with at least 50% maxTOTPAR are shown in *Table 9*. The calculated number of patients in each treatment group with at least 50% maxTOTPAR was derived from 45 actual treatments, using the relationship between mean % maxTOTPAR and percentage > 50% maxTOTPAR. Mean % maxTOTPAR for each study was entered into the equation to the regression to derive the proportion with more than half relief. This proportion was then combined with the number of patients to generate the actual number of patients in each group predicted to have better than 50% relief. Numerical values were rounded up or down to the nearest integer.

Actual mean and proportion with at least 50% maxTOTPAR

The relationship between mean % maxTOTPAR and proportion with at least 50% maxTOTPAR is shown in *Figure 12*. The equation to the regression line was:

Percentage of patients > 50% maxTOTPAR = 1.41 % maxTOTPAR – 14.1 (r^2 = 0.89)

Calculated number of patients with at least 50% maxTOTPAR

The actual and calculated numbers of patients in each group with at least 50% maxTOTPAR are shown in *Table 9*. The equation to the regression line was:

Calculated number of patients > 50% maxTOTPAR = 0.93 actual + 0.93 (r^2 = 0.88)

In 36 of 45 treatments, the agreement between actual and calculated was within two patients; in 42 of 45, agreement was within three patients and in 43 of 45, agreement was within four patients. The two most aberrant results occurred in the same trial.[22]

Simulated actual distribution – mean and proportion with at least 50% maxTOTPAR

A simulated distribution, similar to that of the actual data ('simulated actual' distribution) was used to produce 10,000 simulated treatments. This generated a regression of mean % maxTOTPAR against percentage > 50% maxTOTPAR which was very similar to that obtained for the actual data from 45 treatments:

Percentage of patients > 50% maxTOTPAR = 1.34 mean % maxTOTPAR – 14.1 (r^2 = 0.79)

Simulated actual distribution – calculated numbers > 50% maxTOTPAR

This equation was used to obtain the calculated percentage > 50% maxTOTPAR, which was then regressed against the actual percentage of patients > 50% maxTOTPAR. The equation to the regression line was very similar to that obtained for the actual data from 45 treatments:

Calculated number of patients > 50% maxTOTPAR = 0.82 actual + 1.92 (r^2 = 0.83)

From *Table 10* it can be seen that, using the underlying distribution, the difference between calculated and actual number of patients with at least 50% maxTOTPAR was 0–2 in 90% of the simulated studies and, in 99%, it was in the range 0–3. These results are very similar to those obtained with the actual data and, again, this suggests strongly that provided the underlying actual amalgamated distribution is a reasonable reflection of the assumed 'true' underlying distribution of pain relief, then the conversion technique is accurate and robust.

Normal and uniform distributions

In order to test the effect of different underlying distributions, the process of obtaining the number

TABLE 9 *Studies*

Study	Treatment (N)	Mean % maxTOTPAR	Actual number with > 50% maxTOTPAR	Calculated number with > 50% maxTOTPAR
Evans, *et al.* 1982[6]	paracetamol, 650 mg, + dextropropoxyphene, 65 mg (30)	46.0	18	15
	placebo (30)	29.4	11	8
	zomepirac, 100 mg (30)	38.4	12	12
	zomepirac, 50 mg (30)	49.4	19	17
McQuay, *et al.* 1986[21]	paracetamol, 500 mg (30)	31.0	8	9
	ketorolac, 5 mg (30)	39.5	11	12
	ketorolac, 10 mg (30)	47.0	16	15
	ketorolac, 20 mg (30)	54.0	18	19
	paracetamol, 1000 mg (30)	41.9	12	14
McQuay, *et al.* 1987[8]	aspirin, 650 mg (30)	23.4	7	6
	fluradoline, 150 mg (30)	17.3	7	3
	fluradoline, 300 mg (30)	26.9	8	7
	placebo (30)	10.7	2	0
Porter, *et al.* 1981[5]	bicifadine, 100 mg (19)	12.0	1	1
	bicifadine, 150 mg (20)	17.8	2	2
	placebo (21)	11.9	2	1
	codeine, 60 mg (20)	25.0	4	4
McQuay, *et al.* 1990[9]	bromfenac, 5 mg (30)	25.4	6	6
	bromfenac, 10 mg (30)	38.9	14	12
	bromfenac, 25 mg (30)	46.3	15	15
	placebo (30)	16.9	2	3
	paracetamol, 1000 mg (30)	32.9	10	10
Carroll, *et al.* 1993[20]	bromfenac, 10 mg (23)	58.6	16	16
	bromfenac, 25 mg (21)	46.4	13	11
	buprenorphine, 0.2 mg (22)	21.7	1	4
	buprenorphine, 0.4 mg (24)	35.5	9	9
Bullingham, *et al.* 1981[19]	paracetamol, 1000 mg (30)	51.7	17	18
	paracetamol, 1000 mg + buprenorphine, 1.0 mg (30)	47.8	18	16
	paracetamol, 1000 mg + buprenorphine, 1.5 mg (29)	54.9	15	18
	paracetamol, 1000 mg + buprenorphine, 2.0 mg (30)	50.8	16	17
McQuay, *et al.* 1985[7]	dihydrocodeine, 30 mg (18)	40.6	8	8
	placebo (19)	20.1	4	3
	zomepirac, 100 mg (18)	47.4	11	9
McQuay, *et al.* 1992[24]	paracetamol, 1000 mg, + codeine, 16 mg, + caffeine, 60 mg (30)	39.1	10	12
	ibuprofen, 400 mg + codeine, 25.6 mg (30)	54.0	21	19
McQuay, *et al.* 1993[25]	dihydrocodeine, 30 mg (41)	28.7	9	11
	dihydrocodeine, 60 mg (43)	32.8	13	14
	ibuprofen, 400 mg (40)	60.0	31	28

continued

TABLE 9 contd *Studies*

Study	Treatment (N)	Mean % maxTOTPAR	Actual number with > 50% maxTOTPAR	Calculated number with > 50% maxTOTPAR
McQuay, *et al.* 1989[23]	ibuprofen, 400 mg (23)	44.8	10	11
	ibuprofen, 400 mg + codeine, 20 mg (24)	57.7	15	16
McQuay, *et al.* 1987[22]	aspirin, 500 mg + paracetamol, 500 mg (47)	36.6	13	18
	aspirin, 500 mg, + paracetamol, 500 mg, + codeine, 13.6 mg (48)	34.8	8	17
McQuay, *et al.* unpublished[26]	pethidine, 100 mg (21)	16.1	1	2
	meptazinol, 100 mg (20)	21.8	1	3
	morphine, 15 mg (22)	24.4	4	4

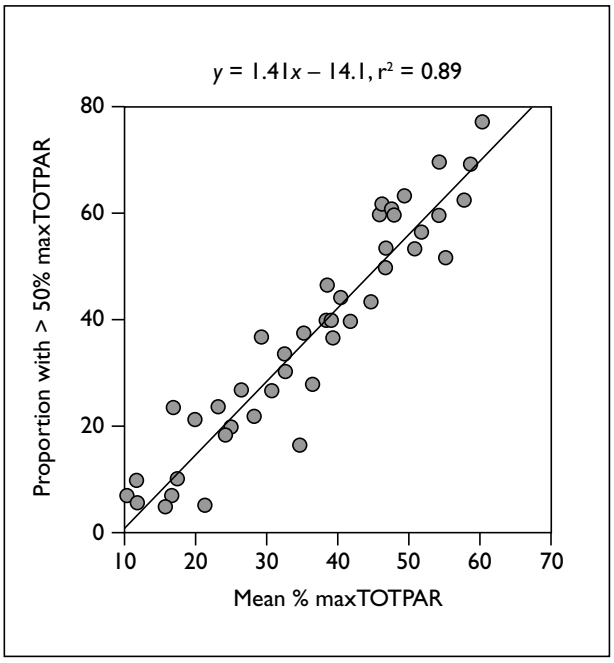


FIGURE 12 *Relationship between mean % maxTOTPAR and proportion with at least 50% maxTOTPAR*

of patients with at least 50% maxTOTPAR was repeated for two further distributions. The results obtained for the 'normal' and 'uniform' distributions *(Table 10)* were less accurate. Even so, these levels of agreement indicate that the conversion technique is robust, even with these gross differences in underlying distribution, and suggest that it will be very robust to the smaller differences likely to be encountered in practice.

NNTs

NNTs were calculated for paracetamol, 1000 mg, zomepirac, 100 mg, bromfenac, 10 mg, bromfenac, 25 mg, dihydrocodeine, 30 mg, ibuprofen, 400 mg, and ibuprofen, 400 mg, plus codeine, 24.6 mg; for these, and for placebo, there was information from at least two trials *(Table 11)*. NNT values derived from the actual and the calculated data, as well as odds ratios and CIs, were very similar or identical.

Single treatment arms from the individual reports were combined to obtain odds ratio estimates with 95% CIs using a fixed effects model and to derive

TABLE 10 *Accuracy of the conversion in actual and simulated treatments*

Difference between actual and calculated numbers	45 actual treatments (%)	Simulated actual distribution (%)	Simulated normal distribution (%)	Simulated uniform distribution (%)
≤ 1	57.7	60.3	21.3	40.7
≤ 2	82.1	90.8	44.4	71.2
≤ 3	93.2	98.8	68.2	89.0
≤ 4	95.4	99.8	87.1	97.1
≤ 5	97.7	100	96.5	99.3
≤ 6	97.7	100	99.4	99.9

TABLE 11 *Numbers-needed-to-treat*

Treatment		Active > 50% maxTOTPAR/Total	NNT (95% CI)	Odds ratio (95% CI)
Dihydrocodeine, 30 mg	Actual	17/59	7.9 (3.9–∞)	2.2 (1.0–4.7)
	Calculated	19/59	4.8 (3.0–13.3)	4.0 (1.8–9.0)
Paracetamol, 1000 mg	Actual	39/90	3.7 (2.6–6.6)	3.9 (2.1–7.1)
	Calculated	42/90	2.9 (2.1–4.3)	6.2 (3.4–11.4)
Zomepirac, 100 mg	Actual	23/48	3.2 (2.1–6.1)	5.5 (2.5–11.7)
	Calculated	21/48	3.1 (2.1–5.8)	7.3 (3.2–16.6)
Bromfenac, 10 mg	Actual	30/53	2.5 (1.8–3.9)	7.4 (3.6–15.1)
	Calculated	28/53	2.4 (1.8–3.7)	9.8 (4.6–20.8)
Bromfenac, 25 mg	Actual	28/51	2.6 (1.9–4.2)	7.0 (3.4–14.6)
	Calculated	26/51	2.5 (1.8–4.1)	9.4 (4.3–20.3)
Ibuprofen, 400 mg	Actual	41/63	2.0 (1.6–2.8)	9.3 (4.9–17.7)
	Calculated	39/63	2.0 (1.6–2.7)	12.0 (6.2–23.4)
Ibuprofen, 400 mg, plus codeine, 24.6 mg	Actual	36/54	2.0 (1.6–2.7)	10.5 (5.3–20.8)
	Calculated	35/54	1.9 (1.5–2.5)	14.6 (7.1–29.6)

NNTs for analgesic effectiveness.[28] At least two identical treatments from different trials were required. Of the 130 patients who received placebo, 21 actually had at least 50% maxTOTPAR and 15 were calculated to have at least 50% maxTOTPAR.

Verification from independent data

Stage 2 methods

Individual patient data from 18 primary RCTs was made available by Grünenthal GmbH, Aachen, Germany, and Robert Wood Johnson Pharmaceutical Research Institute, Spring House, PA, USA (RWJ).

Study protocols for postsurgical pain (including gynaecological procedures) and pain due to the extraction of impacted third molars were essentially identical. Trials were double-blind, single-dose, parallel-group studies; randomisation was by computerised random-number generation, stratified on pretreatment pain. Criteria for patient selection were moderate or severe pain, and that the patient's condition was appropriate for management with a centrally-acting analgesic or paracetamol combined with centrally-acting analgesics. Patients ages ranged from 18 years to 70 years. Patients had to be cooperative, reliable, and motivated, and be able to take oral medication. Exclusion criteria included patients with mild or no pain, those who had taken analgesic drugs within 3 hours of study drug administration, those needing sedatives during the observation period and those with known contraindications or medical conditions which might interfere with observations.

The following drugs were given as single oral doses: placebo (695 evaluable patients); codeine, 60 mg (649); tramadol, 50 mg (409); tramadol, 75 mg (281); tramadol, 100 mg (468); tramadol, 150 mg (279); tramadol, 200 mg (50); aspirin, 650 mg, plus codeine, 60 mg (305); and paracetamol, 650 mg, plus propoxyphene, 100 mg (316).

Patients were given the study drug if they had moderate or severe pain on a four-point categorical scale (0 = no pain, 1 = slight, 2 = moderate, 3 = severe). Thereafter observations were made at 30 minutes, and at 1, 2, 3, 4, 5 and 6 hours after administration. Pain intensity was measured using the same categorical scale, together with a five-point categorical scale of pain relief (0 = no relief, 1 = a little, 2 = some, 3 = a lot, 4 = complete). Time of repeat medication was also recorded, together with a global assessment of therapy (excellent, very good, good, fair or poor) at the final evaluation. At repeat medication, pain relief scores reverted to zero and pain intensity scores to the initial value; adverse event recording, but not pain evaluations, continued after repeat medication.

For each patient the area under the curve of pain relief (categorical scale) against time (TOTPAR)

was calculated for 6 hours after the study drug was given. The percentage of the maximum possible for this summary measure was then calculated for each patient (% maxTOTPAR[10]), mean TOTPAR was calculated for all patients in each treatment arm and the number of patients on each treatment achieving at least 50% maxTOTPAR was noted.

The mean TOTPAR value was then used to calculate the theoretical number of patients with at least 50% maxTOTPAR, using a relationship established in clinical trials of analgesics in Oxford with 1283 patients with 45 treatments (percentage of patients with at least 50% maxTOTPAR = 1.41 x mean % maxTOTPAR – 14.1). Actual and calculated numbers were then compared using unweighted linear regression analysis.

Stage 2 results

Individual patient information was available from over 3400 patients in 85 different treatment arms in nine studies involving dental surgery (mostly third molar extraction) and nine involving general postoperative pain (including gynaecological procedures). Studies involved between 21 and 58 patients in each treatment (mean 40 patients). The distributions of % maxTOTPAR for all active and all placebo patients in these groups are shown in *Figure 13*.

The relationship between actual and calculated numbers of patients with at least 50% maxTOTPAR in each treatment arm is shown in *Figure 14,* and the equation to the regression line for this is compared with 45 treatments from trials in Oxford in *Table 12,* using both the relationship for the actual data and that from a 10,000 treatment simulation.

Of the 85 treatment arms, 80 (94%) were within four patients per treatment and 74 (87%) within three *(Table 13)*. These proportions are comparable to those obtained previously for actual and simulated treatments (see *Table 10)*. Summing the positive and negative differences between actual and calculated numbers of patients with at least 50% maxTOTPAR gave an average difference of 0.30 patients per treatment arm.

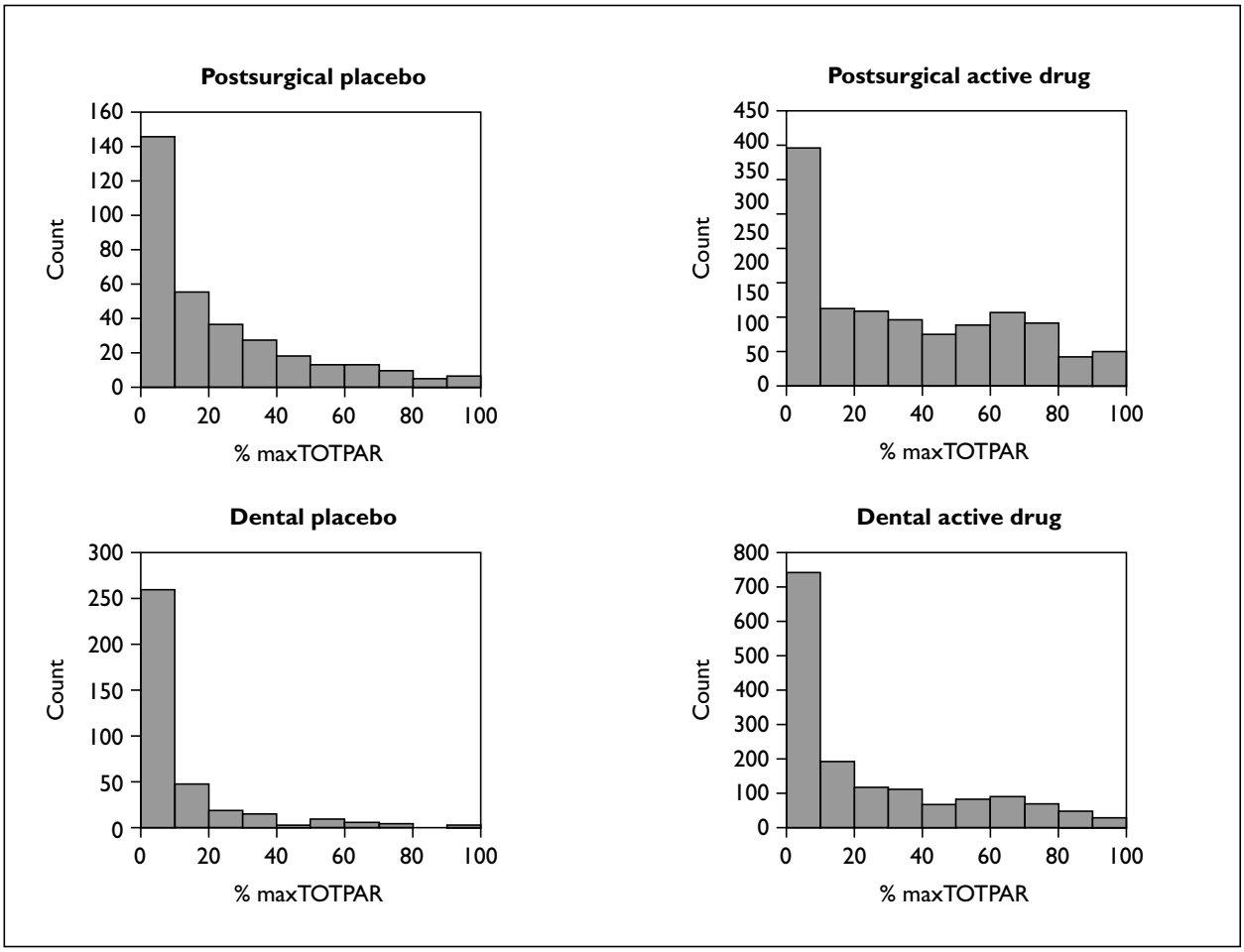


FIGURE 13 *Distributions of % maxTOTPAR for active treatments and placebo in dental and postsurgical pain*

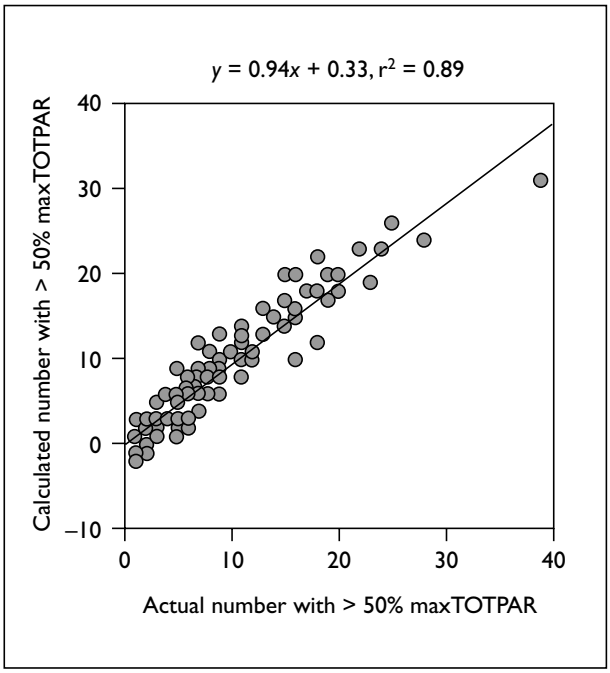


FIGURE 14 *Relationship between actual and calculated number with at least 50% maxTOTPAR for 3400 patients with 85 treatments*

Comparison of (actual – calculated, irrespective of sign) numbers of patients with at least 50% maxTOTPAR as percentages for the 45 actual treatments and 10,000 simulated treatments using the simulated actual, normal and uniform distributions are shown in *Table 13*.[18] Cumulative percentages are shown at different levels of agreement and the final column adds the 85 treatment arms from the RWJ trials.

Combining the 85 treatments in this data set with the earlier 45 treatments[18] produced a new relationship for use in future conversions:

Proportion of patients with > 50% maxTOTPAR = 1.33 x mean % maxTOTPAR – 11.5 ($r^2 = 0.89$)

Use of pain intensity and VAS

Stage 3 methods

Data for the study were from individual patient data from 13 RCTs (1283 patients with 45 treatments, Oxford data[18]) and 18 RCTs (3453 patients with 87 treatments, RWJ data) described in stages 1 and 2.

TABLE 12 *Regression equations for calculated and actual number of patients in each treatment with > 50% maxTOTPAR*

Study	Slope	Intercept	Coefficient of determination (r^2)
45 treatment arms from RCTs in Oxford [1]	0.93	0.93	0.88
45 treatment arms from RCTs in Oxford [2]	0.82	1.92	0.83
85 treatment arms from RWJ RCTs (3453 patients) [3]	0.94	0.33	0.89

Results for solutions to the equation: Calculated = (Actual x slope) + intercept, using the relationship between % > 50% maxTOTPAR and mean % maxTOTPAR derived from:
[1] 45 actual treatments (Oxford RCTs)
[2] a 10,000 treatment arm simulation
[3] 85 actual treatments (RWJ RCTs)

TABLE 13 *Accuracy of the conversion in actual and simulated treatments*

Difference between actual and calculated numbers	45 actual treatments (Oxford) (%)	Simulated actual distribution (%)	Simulated normal distribution (%)	Simulated uniform distribution (%)	85 RWJ actual treatments (%)
≤ 1	57.7	60.3	21.3	40.7	50.6
≤ 2	82.1	90.8	44.4	71.2	70.6
≤ 3	93.2	98.8	68.2	89.0	87.1
≤ 4	95.4	99.8	87.1	97.1	94.1
≤ 5	97.7	100	96.5	99.3	96.6
≤ 6	97.7	100	99.4	99.9	98.8

For each patient the SPID was calculated for categorical pain intensity, and the equivalent VAS SPID. For each individual patient the 4-hour or 6-hour SPID was divided by the maximum possible SPID; for example, a patient with a SPID of 6 and initial pain intensity of 3 would have a theoretical maximum SPID of 18, and the % maxSPID would be 33%. The area under the curve of pain relief against time was calculated for the categorical (TOTPAR) and VAS TOTPAR scales. The percentage of the maximum possible for each summary measure was then calculated for each patient.[10] Rules for calculation included that in the event of repeat medication within 6 hours, pain relief scores reverted to zero and pain intensity scores to their initial value. The mean summary measure for all patients in each treatment arm was calculated. The number of patients on each treatment achieving at least 50% maxTOTPAR was noted.

The relationship between the mean % maxSPID, % maxVAS SPID and % maxVAS TOTPAR and the actual number of patients with at least 50% maxTOTPAR was examined by linear regression analysis. Using the equation to the regression line, the calculated number of patients with at least 50% maxTOTPAR was then compared with the actual number using unweighted linear regression analysis.

Stage 3 results

Individual patient scores for categorical pain intensity and VAS pain intensity and pain relief were asymmetrically distributed, much as was seen for TOTPAR.

Categorical pain intensity scale

Data were available from 132 treatments with 4713 patients. Individual patient distribution of % maxSPID was asymmetric *(Figure 15)*. Linear regression analysis performed for the Oxford and RWJ data sets separately showed similar relationships, so the data sets were combined for all 132 treatments.

Results from Oxford and RWJ data sets and the combined data for the regression of number of patients per treatment with at least 50% maxTOTPAR against mean %maxSPID (with 95% CI).

For all 132 treatments the regression line was:

Percentage with > 50% maxTOTPAR
= 1.36 mean % maxSPID – 2.3 ($r^2 = 0.85$)

There was good agreement between the actual number of patients with at least 50% maxTOTPAR in each treatment arm and the calculated number

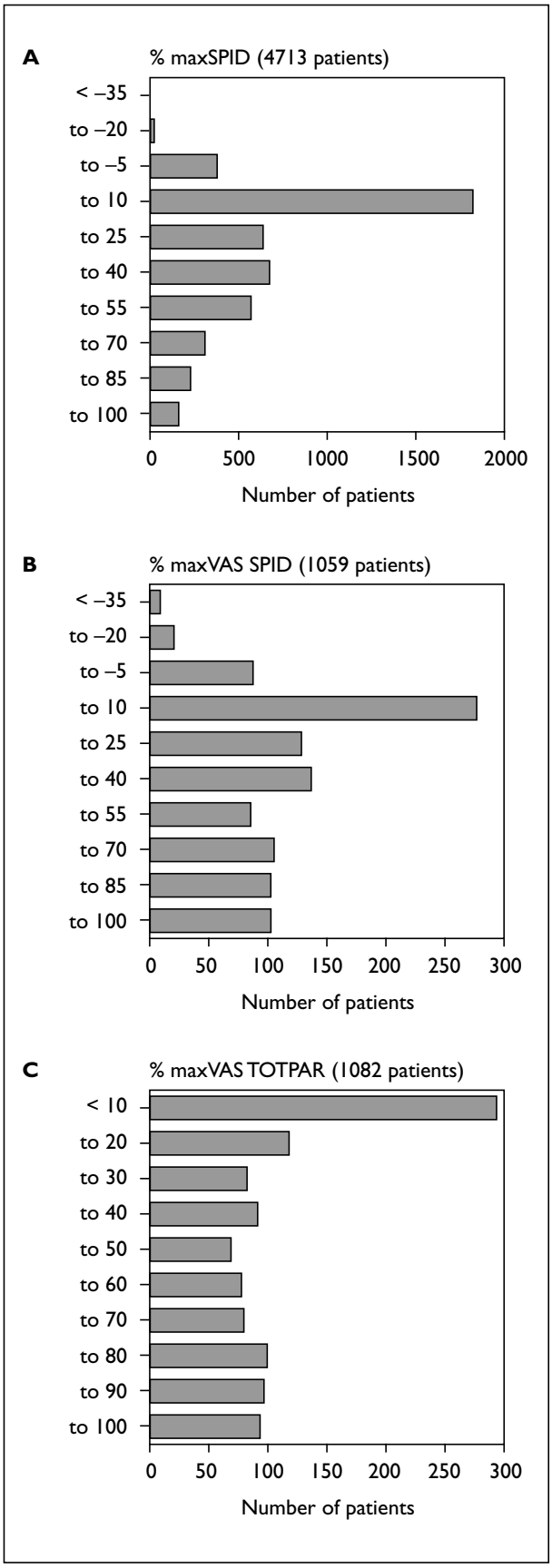


FIGURE 15 *Distribution of individual patient scores for categorical pain intensity and VAS pain intensity and pain relief scores*

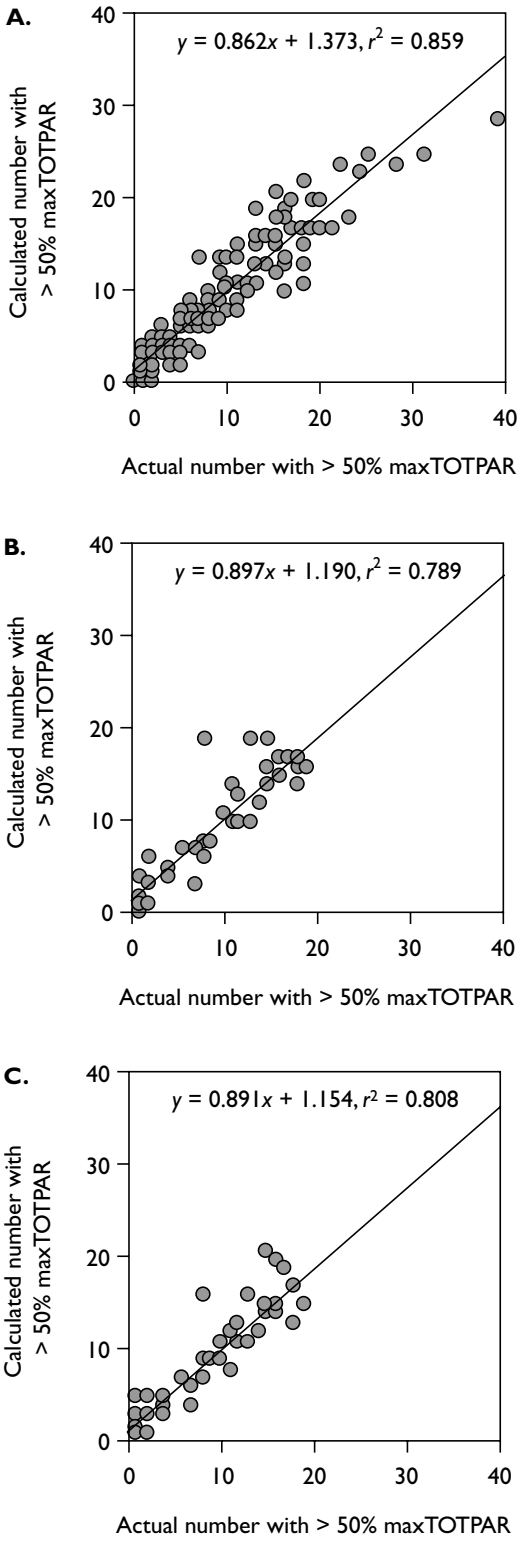


FIGURE 16 *Correlation of actual and calculated numbers of patients with at least 50% maxTOTPAR in each treatment for calculations using categorical pain intensity and VAS pain intensity and relief scores*

using the relationship derived with % maxSPID *(Figure 16):*

Calculated number with > 50% maxTOTPAR = 0.86 actual + 1.37 (r^2 = 0.86)

For 92% of treatments the actual and the calculated numbers with at least 50% maxTOTPAR were within four patients per treatment. Agreement (actual – calculated) was normally distributed around zero *(Figure 17)*. Summing the positive and negative differences between actual and calculated numbers of patients with at least 50% maxTOTPAR gave an average difference of –0.03 patients per treatment arm.

VAS pain intensity

Data were available from 40 treatments within the Oxford data set with 1059 patients. Individual patient distribution of % maxVAS SPID was asymmetric *(Figure 15B)*. The regression line between percentage with > 50% maxTOTPAR and mean % maxVAS SPID was given by:

Percentage with > 50% maxTOTPAR = 1.18 mean % maxVAS SPID – 2.2 (r^2 = 0.87)

There was good agreement between the actual number of patients with at least 50% maxTOTPAR in each treatment arm and the calculated number using the relationship derived from % maxVAS SPID *(Figure 16B):*

Calculated number with > 50% maxTOTPAR = 0.90 actual + 1.19 (r^2 = 0.79)

For 95% of treatments the actual and the calculated numbers with at least 50% maxTOTPAR were within four patients per treatment. Summing the positive and negative differences between actual and calculated numbers of patients with at least 50% maxTOTPAR gave an average difference of –0.23 patients per treatment arm.

VAS pain relief

Data were available from 40 treatments with 1082 patients. Individual patient distribution of % maxVAS TOTPAR was asymmetric *(Figure 15C)*. The regression line between percentage with > 50% maxTOTPAR and mean % maxVAS TOTPAR was given by:

Percentage > 50% maxTOTPAR = 1.15 mean % maxVAS TOTPAR – 8.51 (r^2 = 0.81)

There was good agreement between the actual number of patients with at least 50% maxTOTPAR in each treatment arm and the calculated number

using the relationship derived from % maxVAS TOTPAR *(Figure 16C):*

Calculated number with > 50% maxTOTPAR = 0.89 actual + 1.15 ($r^2 = 0.81$)

For 95% of treatments the actual and the calculated numbers with at least 50% maxTOTPAR were within four patients per treatment. Summing the positive and negative differences between actual and calculated numbers of patients with at least 50% maxTOTPAR gave an average difference of –0.11 patients per treatment arm.

Overall comments

For SPID it was possible to use the gold standard of verification by independent data sets. Regressing the percentage > 50% maxTOTPAR against mean % maxSPID independently for Oxford and RWJ data sets produced very similar results *(Table 14).* Using the combined regression analysis, there was excellent agreement between actual and calculated numbers of patients with at least 50% maxTOTPAR in each treatment *(Figure 16A),* and the sum of the difference over all 132 treatments was –0.03 patients per treatment, with the differences distributed normally around zero *(Figure 17).* This is firm evidence for the reliability of the conversion method.

Only 40 treatments from the Oxford data set were available for calculating relationships between patients with at least 50% maxTOTPAR and mean % maxVAS SPID and mean % maxVAS TOTPAR. Despite this, the agreement between actual and calculated numbers with at least 50% maxTOTPAR was good *(Figure 16B, C,* and *Table 15),* so that over the 40 treatments the sum of actual minus calculated was less than a quarter of a patient per treatment arm using either measure.

Although no independent verification was possible for VAS, the similarity of the results to those independently verified for TOTPAR[18,30] and SPID supports the approach of using mean data from previously published reports to derive dichotomous data for meta-analysis.[18]

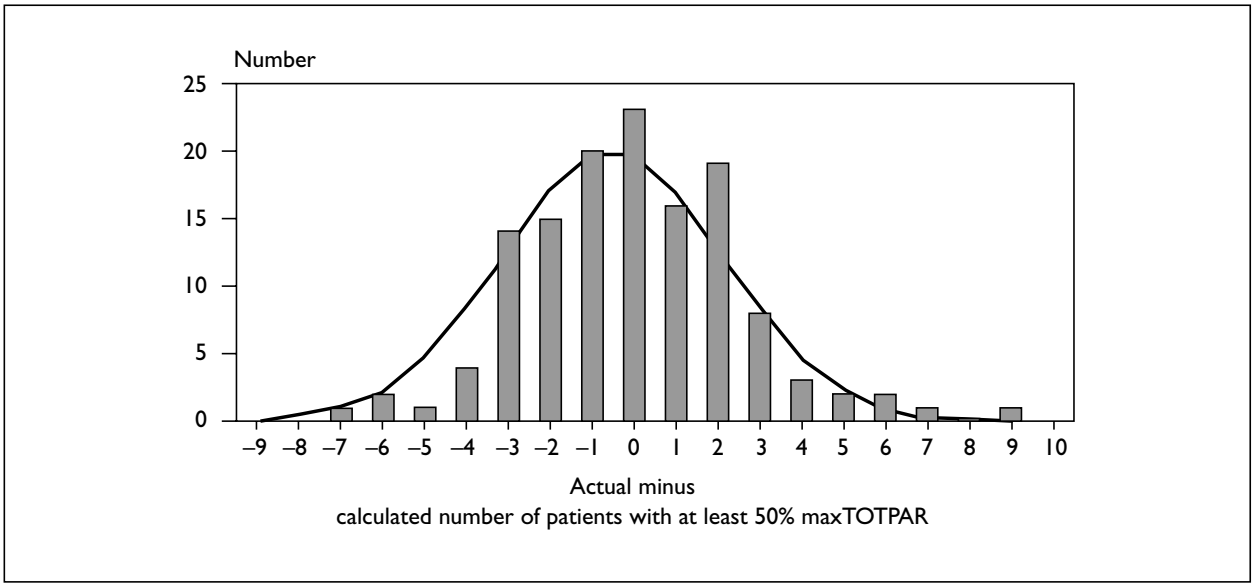


FIGURE 17 *Distribution of actual – calculated number of patients with at least 50% maxTOTPAR in each treatment using SPID data*

TABLE 14 *Summary report on SPID calculations for 132 treatments*

Results from Oxford and RWJ data sets and the combined data for the regression of number of patients per treatment with > 50% maxTOTPAR against mean % maxSPID (with 95% CI).

Data set	N	Intercept (95% CI)	Slope (95% CI)	Coefficient of determination
Oxford	45	–2.3 (–8.3, 3.6)	1.44 (1.24, 1.64)	0.83
RWJ	87	–1.7 (–4.3, 1.0)	1.27 (1.16, 1.38)	0.86
Combined	132	–2.3 (–4.9, 0.2)	1.36 (1.26, 1.45)	0.85

Comment

There is an asymmetric distribution of summary values of pain relief in clinical trials of analgesics using standard trial methods. Using mean values to describe these summary values is inappropriate and may result in erroneous conclusions.[27] To use information from RCTs of analgesic drugs reporting mean data, conversion to some form amenable to meta-analysis is necessary – and, preferably, some dichotomous measurement. The alternative may be to discard the many thousands of studies of analgesic interventions in the literature.

Some possible methods of conversion have been subjected to the gold standard of verification by an independent data set. There were many patients, in many studies, with different clinical settings, using placebo and several different active analgesics. The result – the relationship between the calculated and actual number of patients with at least 50% maxTOTPAR – was essentially the same as that obtained originally using the relationship for the actual data and from a 10,000 treatment arm simulation. Verification was also possible for SPID, but not for VAS, though there is no obvious reason to suspect that conversions explored here should not be accurate.

From the categorical pain relief scale and its summary TOTPAR measure, dichotomous data (the proportion of patients achieving at least 50% of % maxTOTPAR and the corollary, those not achieving 50% relief) can now be derived with some confidence. Categorical pain relief data can also be used with confidence.

Other data may be used as it becomes available to further validate these relationships *(Table 16),* based on a wide variety of acute pain conditions with different analgesics, including simple analgesics, NSAIDs, combinations and sublingual and intramuscular opiates. The only caution is that the validity of these relationships has been demonstrated only in short-term single-dose studies in acute pain models.

References

1. Wall PD. The placebo and the placebo response. In: Wall PD, Melzack R, editors. Textbook of pain. 3rd ed. Edinburgh: Churchill-Livingstone, 1994:1297–1308.
2. Wall PD. The placebo effect: an unpopular topic. *Pain* 1992;**51**:1–3.
3. Beecher HK. The powerful placebo. *JAMA* 1955;**159**:1602–6.

TABLE 15 *Accuracy of the conversion in actual and simulated treatments*

		Actual minus calculated % of patients with > 50% maxTOTPAR for treatment arms: agreement to within –					
Data sets	**Scale**	**≤ 1**	**≤ 2**	**≤ 3**	**≤ 4**	**≤ 5**	**≤ 6**
45 Oxford	TOTPAR	58	82	93	95	98	98
85 RWJ	TOTPAR	52	75	85	94	96	100
132 Oxford and RWJ combined	SPID	45	70	87	92	95	98
40 Oxford	VAS SPID	65	75	85	95	95	98
40 Oxford	VAS TOTPAR	65	73	85	95	98	98

Comparisons of actual minus calculated (irrespective of sign) percentages of patients with at least 50% maxTOTPAR as cumulative percentages for the 45 Oxford treatments using TOTPAR and 85 RWJ treatments using TOTPAR.[18,30] Cumulative percentages are shown at different levels of agreement. The final three rows show comparisons using SPID, VAS SPID and VAS TOTPAR as basis of calculations in Oxford and RWJ data sets.

TABLE 16 *Summary of formulae to derive proportion of patients achieving at least 50% pain relief from mean data using different outcome measures*

Outcome measure	**Proportion of patients achieving at least 50% PR**	**Formula**
Categorical PR		1.33 x mean % maxTOTPAR – 11.5
Categorical PI		1.36 mean % maxSPID – 2.3
VAS PR		1.15 mean % maxVAS TOTPAR – 8.51
VAS PI		1.18 mean % maxVAS SPID – 2.2

4. Evans FJ. The placebo response in pain reduction. *Adv Neurol* 1974;**4**:289–96.
5. Porter EJB, Rolfe M, McQuay HJ, Moore RA, Bullingham RES. Single dose comparison of bicifadine, codeine and placebo in postoperative pain. *Curr Ther Res* 1981;**30**:156–60.
6. Evans PJ, McQuay HJ, Rolfe M, O'Sullivan G, Bullingham RE, Moore RA. Zomepirac, placebo and paracetamol/dextropropoxyphene combination compared in orthopaedic postoperative pain. *Br J Anaesth* 1982;**54**:927–33.
7. McQuay HJ, Bullingham RE, Moore RA, Carroll D, Evans PJ, O'Sullivan G, *et al.* Zomepirac, dihydrocodeine and placebo compared in postoperative pain after day-case surgery. The relationship between the effects of single and multiple doses. *Br J Anaesth* 1985;**57**:412–19.
8. McQuay HJ, Carroll D, Poppleton P, Summerfield RJ, Moore RA. Fluradoline and aspirin for orthopedic postoperative pain. *Clin Pharmacol Ther* 1987;**41**:531–6.
9. McQuay HJ, Carroll D, Frankland T, Harvey M, Moore A. Bromfenac, acetaminophen, and placebo in orthopedic postoperative pain. *Clin Pharmacol Ther* 1990;**47**:760–6.
10. Cooper SA. Single-dose analgesic studies: the upside and downside of assay sensitivity. *Adv Pain Res Ther* 1991;**18**:117–24.
11. Gracely RH, Dubner R, Deeter WR, Wolskee PJ. Clinicians' expectations influence placebo analgesia. *Lancet* 1985;**i**:43.
12. Wall PD. Pain and the placebo response. Experimental and theoretical studies of consciousness. CIBA Foundation Symposium 174. Chichester: Wiley, 1993:187–216.
13. Bergmann J-F, Chassany O, Gandiol J, Deblois P, Kanis JA, Segrestaa J, *et al.* A randomised clinical trial of the effect of informed consent on the analgesic activity of placebo and naproxen in cancer pain. *Clin Trials Meta-Analysis* 1994;**29**:41–7.
14. Shapiro AP, Myers T, Reiser MF, Ferris EB. Comparison of blood pressure response to veriloid and to the doctor. *Psychosom Med* 1954;**16**:478–88.
15. Gøtzsche PC. Meta-analysis of NSAIDs: contribution of drugs, doses, trial designs, and meta-analytic techniques. *Scand J Rheumatol* 1993;**22**:255–60.
16. Moertel CG, Ahmann DL, Taylor WF, Schwartau N. Relief of pain by oral medications. *JAMA* 1974;**229**:55–9.
17. Jadad AR. Meta-analysis of randomised clinical trials in pain relief [DPhil thesis]. Oxford: University of Oxford, 1994.
18. Moore A, McQuay H, Gavaghan D. Deriving dichotomous outcome measures from continuous data in randomised controlled trials of analgesics. *Pain* 1996;**66**:229–37.
19. Bullingham RE, McQuay HJ, Moore RA, Weir L. An oral buprenorphine and paracetamol combination compared with paracetamol alone: a single dose double-blind postoperative study. *Br J Clin Pharmacol* 1981;**12**:863–7.
20. Carroll D, Frankland T, Nagle C, McQuay H. Oral bromfenac 10 and 25 mg compared with sublingual buprenorphine 0.2 and 0.4 mg for postoperative pain relief. *Br J Anaesth* 1993;**71**:814–17.
21. McQuay HJ, Poppleton P, Carroll D, Summerfield RJ, Bullingham RE, Moore RA. Ketorolac and acetaminophen for orthopedic postoperative pain. *Clin Pharmacol Ther* 1986;**39**:89–93.
22. McQuay HJ, Carroll D, Watts PG, Juniper RP, Moore RA. Does adding small doses of codeine increase pain relief after third molar surgery? *Clin J Pain* 1987;**2**:197–201.
23. McQuay HJ, Carroll D, Watts PG, Juniper RP, Moore RA. Codeine 20 mg increases pain relief from ibuprofen 400 mg after third molar surgery. A repeat-dosing comparison of ibuprofen and an ibuprofen-codeine combination. *Pain* 1989;**37**:7–13. [Erratum: *Pain* 1989;**39**:373].
24. McQuay HJ, Carroll D, Guest P, Juniper RP, Moore RA. A multiple dose comparison of combinations of ibuprofen and codeine and paracetamol, codeine and caffeine after third molar surgery. *Anaesthesia* 1992;**47**:672–7.
25. McQuay HJ, Carroll D, Guest PG, Robson S, Wiffen PJ, Juniper RP. A multiple dose comparison of ibuprofen and dihydrocodeine after third molar surgery. *Br J Oral Maxillofac Surg* 1993;**31**:95–100.
26. McQuay HJ, Moore RA, Poppleton P, Uppington J, Bullingham RES, Lammer P. Postoperative single-dose comparison of intramuscular meptazinol, morphine and pethidine, compared in the same patients with intravenous meptazinol, morphine and pethidine given by PCA. Unpublished.
27. McQuay H, Carroll D, Moore A. Variation in the placebo effect in randomised controlled trials of analgesics: all is as blind as it seems. *Pain* 1996;**64**:331–5.
28. Cook RJ, Sackett DL. The number needed to treat: a clinically useful measure of treatment effect. *BMJ* 1995;**310**:452–4.
29. Dudewicz EJ, Mishra SM. Modern mathematical statistics. New York: Wiley, 1988.
30. Moore A, McQuay H, Gavaghan D. Deriving dichotomous outcome measures from continuous data in randomised controlled trials of analgesics: verification from independent data. *Pain* 1997;**69**:127–30.
31. Silverstein FE, Graham DY, Senior JR, Davies HW, Struthers BJ, Bittman RM, *et al.* Misoprostol reduces serious gastrointestinal complications in patients with rheumatoid arthritis receiving nonsteroidal anti-inflammatory drugs. *Ann Intern Med* 1995;**123**:241–9.

Chapter 6

Combining data and interpreting results

As professionals we want to use the best treatments and, as patients, to be given them. Knowing that an intervention works (or does not work) is fundamental to clinical decision-making.

When is the evidence strong enough to justify changing practice? Some of the decisions we make are based on individual studies, often on small numbers of patients, which, given the random play of chance, may lead to incorrect decisions. Systematic reviews identify and review all the relevant studies, and are more likely to give a reliable answer. Explicit methods and quality standards are used to reduce bias. The results are the closest we are likely to get to the truth in the current state of knowledge.

The following questions should be answered by a systematic review.

- How well does an intervention work (compared with placebo, no treatment or other interventions in current use) – or can it be forgotten?
- Is it safe?
- Will it work and be safe for the patients in my practice?

Clinicians live in the real world and are busy people; they need to synthesise their knowledge of a particular patient in their practice, their experience and expertise, and the best external evidence from systematic review. They can then be reasonably sure that they are doing their best. However, the product of a systematic review, particularly a meta-analysis, is often some sort of statistical output, which is not usually readily interpretable or usable in day-to-day clinical practice. A common currency to help make the best treatment decision for a particular patient is needed. That common currency is, in the authors' opinion, the NNT.

Quality control

Systematic reviews of inadequate quality may be worse than none, because faulty decisions may be made with unjustified confidence. Quality control in the systematic review process, from literature searching onwards, is vital. How the quality of a systematic review may be judged is encapsulated in the following questions[1] (these are explained in more detail in chapter 7).

- Were the question(s) and methods clearly stated?
- Were the search methods used to locate relevant studies comprehensive?
- Were explicit methods used to determine which articles to include in the review?
- Was the methodological quality of the primary studies assessed?
- Were the selection and assessment of the primary studies reproducible and free from bias?
- Were differences in individual study results explained adequately?
- Were the results of the primary studies combined appropriately?
- Were the reviewers' conclusions supported by the data cited?

When systematic reviews use data from different numbers of papers (see the paper by Vander Stichele and colleagues[2] for an excellent discussion of the eligibility criteria for trials of head lice infection), reasons should be sought. Reviews can use criteria that exclude information important to individual clinicians, or may include studies with inadequate trial design. Inclusion and exclusion criteria must be read critically to see if they make sense in the particular clinical circumstance.

Outcome measures chosen for data extraction should also be sensible. This is not usually a problem but, again, it is a part of the method that needs to be read carefully to see if the outcome measure extracted appears appropriate. The reviewer may have used all the available information, with any problems being due to the original trials, but outcome measures are a determinant of the clinical utility of the review. Examples, in the antibiotic treatment of *Helicobacter pylori* infection and peptic ulcer, would be outcome measures of short-term bacterial kill rates and long-term remission.

Therapeutic interventions: which study architectures are admissible?

The gold standard for a systematic review of therapeutic efficacy is that the eligible studies should be RCTs. If trials are not randomised, estimates of treatment effect may be exaggerated by up to 40% (see *Table 17*).[3] In a systematic review of TENS in postoperative pain, 17 reports on

TABLE 17 *Systematic reviews should eliminate bias*

Feature	Overestimate of treatment effect (%)
Randomisation	40
Double-blind	17
Duplicates	20
Small trials	30

786 patients could be regarded unequivocally as RCTs in acute postoperative pain. Of these 17 RCTs, 15 demonstrated no benefit of TENS over placebo. A total of 19 reports had pain outcomes but were not RCTs; in 17 of these, TENS was considered by their authors to have had a positive analgesic effect.[4] When appropriate, and particularly with subjective outcomes, the gold standard for an efficacy systematic review is studies which are both randomised and double-blind. The therapeutic effect may be exaggerated by up to 20% in trials with deficient blinding.[3]

Not all data can be combined in a meta-analysis: qualitative systematic reviews

It is often not possible or sensible to combine or pool data, and this results in a qualitative rather than a quantitative systematic review. Combining data is not possible if there is no quantitative information in the component trials of the review. Combining data may not be sensible if different clinical outcomes were used or the patients were followed-up for different lengths of time. Combining continuous rather than dichotomous data may be difficult. Even if dichotomous data is measured and presented, if the trials are otherwise of poor quality[5] it may not be sensible to combine the data.

Making decisions from qualitative systematic reviews

Making decisions about whether or not a therapy works from such a qualitative systematic review may look easy. In the example above, 15 of 17 RCTs of TENS in acute pain showed no benefit compared with controls. The thinking clinician will catch the Bayesian drift – that TENS in acute pain is not effective. The problem with such simple vote-counting is that it may mislead. It ignores the sample size of the constituent studies, the magnitude of the effect in the studies and the validity of their design even though they were randomised.

Combining data: quantitative systematic reviews

There are two parts to the question, 'Does it work?' – how does it compare with placebo and how does it compare with other therapies? Whichever comparison is considered, the three stages of examining a review are a L'Abbé plot *(Figure 18),* statistical testing (odds ratio or relative risk), and a clinical significance measure such as NNT.

L'Abbé plots[6]

For therapies, a first stage is to look at a simple scatter plot, which can yield a surprisingly comprehensive qualitative view of the data. Even if the review does not show the data in this way, they can be extracted from information on individual trials presented in the review tables. Data extracted from three different systematic reviews of treatments for painful diabetic neuropathy are shown in *Figure 19*.[7–9] Each point on the graph is the result of a single trial, and what happens with the intervention in question – the experimental event rate (EER) – is plotted against the event rate in the controls – the control event rate (CER).

Trials in which the experimental treatment proves better than the control (EER > CER) will be in the upper left of the plot, between the *y* axis and the line of equality *(Figure 20)*. All three interventions presented in *Figure 19* were effective but the figure does not indicate how effective. If experimental is no better than control then the point will fall on the line of equality (EER = CER), and if control is better than experimental, then the point will be in the lower right of the plot, between the *x* axis and the line of equality (EER < CER).

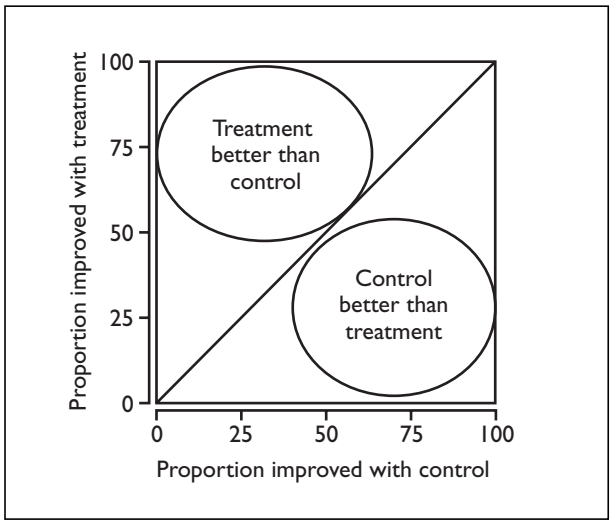


FIGURE 18 *L'Abbé plot for treatment*

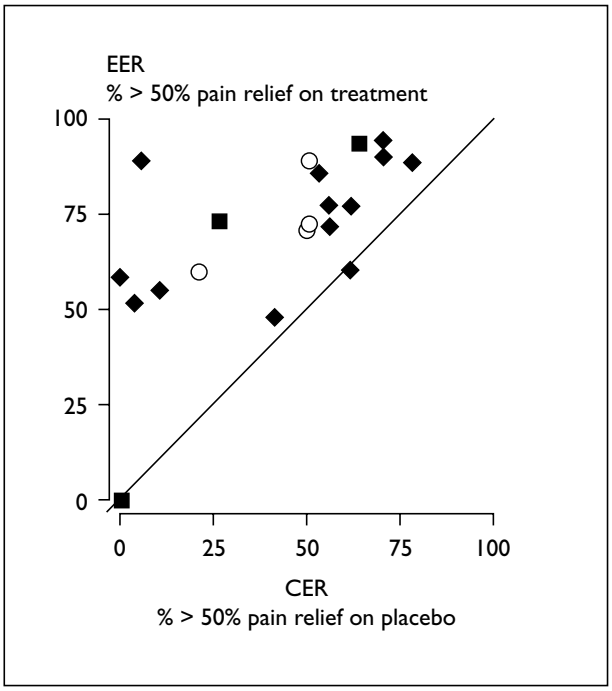


FIGURE 19 *L'Abbé plot of EER against CER (■, RCTs of anticonvulsants; ◆, antidepressants; ○, topical capsaicin in diabetic neuropathy[7–9])*

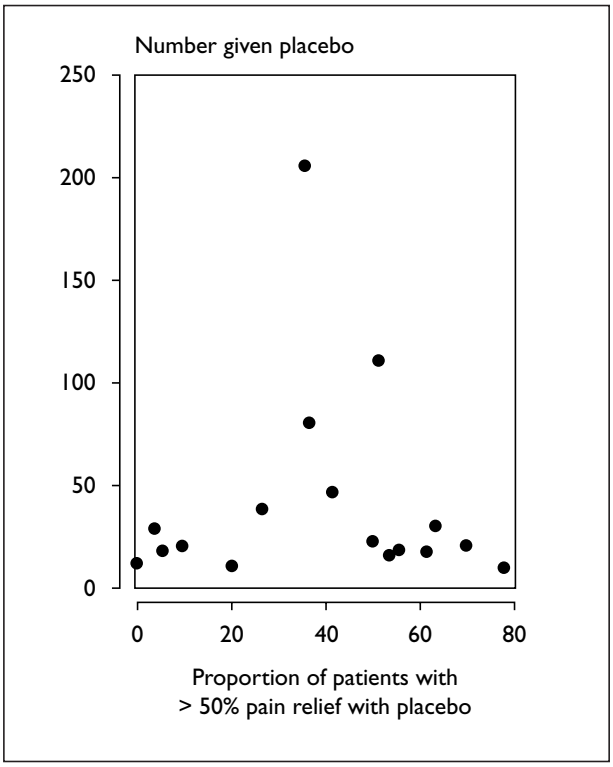


FIGURE 20 *Relationship between placebo response and trial size for pharmacological interventions in diabetic neuropathy*

Visual inspection gives a quick and easy indication of the level of agreement between trials. Heterogeneity is often assumed to be due to variation in the EER, the effect of the intervention. Variation in CER can also be a source of heterogeneity (see *Figure 19);* in this case, the controls were all matched with placebo in a relatively homogenous chronic condition with treatment over a period ranging from several weeks to several months.

L'Abbé plots are not yet widely used. They have several benefits:

- the simple visual presentation is easy to assimilate
- they lead to consideration of the reasons for such wide variation in (especially) placebo responses, and about other factors in the overall package of care that can contribute to effectiveness
- the need for placebo controls is explained if ethical issues about future trials arise
- overly good or bad results for an intervention are viewed with scepticism in a single trial where the major influence may be how good or bad the response was with placebo.

Variation in control (placebo) response rates

The large variation in CER (from 0% to 80%) is not unusual. Similar variation was seen in trials of anti-emetics in postoperative vomiting[10] and, in six trials of prophylactic natural surfactant for preterm infants, the CER for bronchopulmonary dysplasia was 24–69%.[11] Such variation would not be expected in other circumstances, such as in the use of antimicrobial agents. *H. pylori* eradication rates with short-term use of ulcer healing drugs were 0–17% in 11 RCTs (with 10/11 being below 10%).[12]

The reason for large variations in event rates with placebo may have something to do with trial design and population. The overwhelming reason for large variations in placebo rates in pain studies (and probably studies in other clinical conditions) is the relatively small group sizes in trials. Group sizes are chosen to produce statistical significance through power calculations – for pain studies the usual size is 30–40 patients for a 30% difference between placebo and active.

An individual patient can have no pain relief or 100% pain relief. Random selection of patients can therefore produce groups with low or high placebo response rates, or with a rate in between. Ongoing mathematical modelling based on individual patient data shows that, while group sizes of up to 50 patients are likely to show a statistical difference between 80% and 90% of the time, to generate a close approximation to the 'true' clinical impact of a therapy requires as many as 500 patients per group (or more than 1000 patients in a trial). This is part of the rationale of systematic reviews.

Examples of the way group size can be a source of variation are important in understanding how pooling of information in pain trials can be of help. One example, given in *Figure 20,* shows trials in diabetic neuropathy in which the proportion of patients given placebo is plotted against the number given placebo.

A similar pattern of an inverted 'V' is also seen in topical NSAID trials, and indicates that almost all the variability in placebo responses occurs in trials of small size. In rheumatoid arthritis, Gøtzsche[13] found a similar variability in estimates of change in erythrocyte sedimentation rate (ESR) and joint size by sample size.

The lessons are that information from individual trials of small size should be treated with circumspection in pain and probably in other therapeutic areas, and that the variation in outcomes seen in trials of small size is probably an artefact, especially in the absence of any Bayesian drift.

Indirect comparisons

Indirect comparisons of the efficacy of different interventions, for example, by trying to compare treatments which have each been compared with placebo rather than with each other, may not be viable if the CERs are dissimilar. *Post-hoc* approaches, taking all the trials, then using only those which have a low or a high CER, are frowned on, although using particular clinical settings and anticipating less spread of the CER may be more acceptable.[14] In some circumstances, for instance, in prophylaxis for nausea and vomiting, particular CER spreads may be determinants of trial validity.[15]

In most pain studies neither of these apply.

Statistical significance

Odds ratios

When it is legitimate and feasible to combine data, the odds ratio and the relative risk are the accepted statistical tests to show that the intervention works significantly better than the comparator. As more use is made of systematic reviews to compare therapies, clinicians need to understand these clinical epidemiological tools, which present the results in an unfamiliar way.

The odds ratios for the trials of antidepressants in diabetic neuropathy mentioned above are shown in *Figure 21.* Some of the component trials did not show statistical significance; the lower 95% CI of the odds ratio was less than 1. Conversely, other trials and the combined analysis did show statistical significance, with the lower 95% CI being greater than 1, meaning that in 19 cases out of 20 the 'true' value will be greater than 1.

The odds ratio can give a distorted impression when analyses are conducted on subgroups which differ substantially in baseline risk.[14] When CERs are high (certainly when they are above 50%), odds ratios should be interpreted with caution.

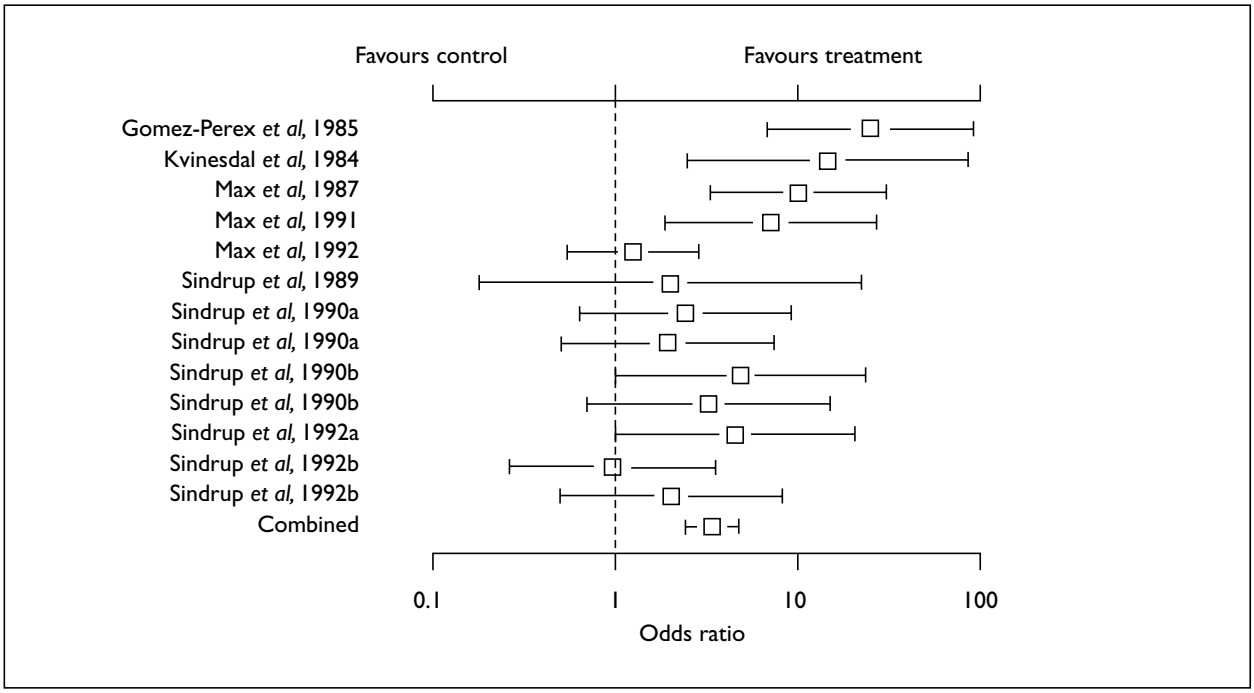


FIGURE 21 *Odds ratios for antidepressants in diabetic neuropathy*[9]

Relative risk or benefit

The fact that it is the odds ratio rather than relative risk or benefit which is used as the test of statistical significance in systematic reviews seems to be due to custom and practice rather than to any inherent intellectual advantage.[14] Relative risk or benefit may be better than an odds ratio because it is more robust in situations where the CER is high.[16] With event rates above 10%, relative risk produces more conservative figures.[17]

In the following chapters, both relative benefit and relative risk (of harm) are used, despite the uncertainty and disagreement between statisticians and reviewers. In all cases, the actual numbers are given so that, when the dust has settled, re-calculations can be made according to the prevailing opinion.

Heterogeneity

Clinicians making decisions on the basis of systematic reviews need to be confident that apples are not being compared with oranges. The L'Abbé plot is a qualitative defence against this spectre. Statistical testing provides a quantitative rampart, and is available as standard software.* Unfortunately, all of these tests lack power, so while a positive test for heterogeneity suggests mixed fruits are being compared, a negative test does not provide complete reassurance of no heterogeneity.

Heterogeneity will also appear to occur because of variations in CERs and EERs caused by the random play of chance in trials of small size. Generally trials with fewer than ten patients per group have been omitted in reviews in this report but considerable variability will occur in groups of less than 50 patients.

How well does an intervention work?

While odds ratios and relative risks can show that an intervention works well compared with a control, they are of limited help in showing how well the intervention works – the size of the effect or its clinical significance.

Effect size

The classical method of estimating effect size is to use the standardised mean difference.[18] The advantage of this approach is that it can be used to compare the efficacy of different interventions measured on continuous rather than dichotomous scales, and even using different outcome measures. The *z* score output is in standard deviation units and is thus scale-free.

The disadvantage of effect size is that it is not intuitive for clinicians.

Number-needed-to-treat

The NNT concept is proving to be a very effective alternative as the measure of clinical significance from quantitative systematic reviews. It has the crucial advantage of applicability to clinical practice, and shows the effort required to achieve a particular therapeutic target. The NNT is given by the equation:

$$NNT = \frac{1}{(IMP_{act}/TOT_{act}) - (IMP_{con}/TOT_{con})}$$

where:

IMP_{act} = number of patients given active treatment who achieve the target

TOT_{act} = total number of patients given active treatment

IMP_{con} = number of patients given a control treatment who achieve the target

TOT_{con} = total number of patients given control treatment

NNT is also 1 divided by the proportion obtaining a particular effect with treatment **minus** the proportion obtaining the same effect with control, when those proportions are expressed as a fraction. Because we have just described the absolute risk reduction, so NNT is also the reciprocal of the absolute risk reduction.

Treatment-specific

NNT is treatment-specific. It describes the **difference** between active treatment and control. The threshold used to calculate NNT can vary but NNT is likely to be relatively unchanged because a change in threshold changes results for both active and control.

For example, in an individual patient data meta-analysis of postoperative pain relief, NNTs compared with placebo were calculated for paracetamol, 650 mg, plus propoxyphene, 100 mg, at between 20 and 80% relief of pain (see *Figure 22).* With placebo, the proportion of patients achieving a particular level of pain relief fell quickly as the

* Cochrane Collaboration: Review Manager Software (RevMan) 1996.

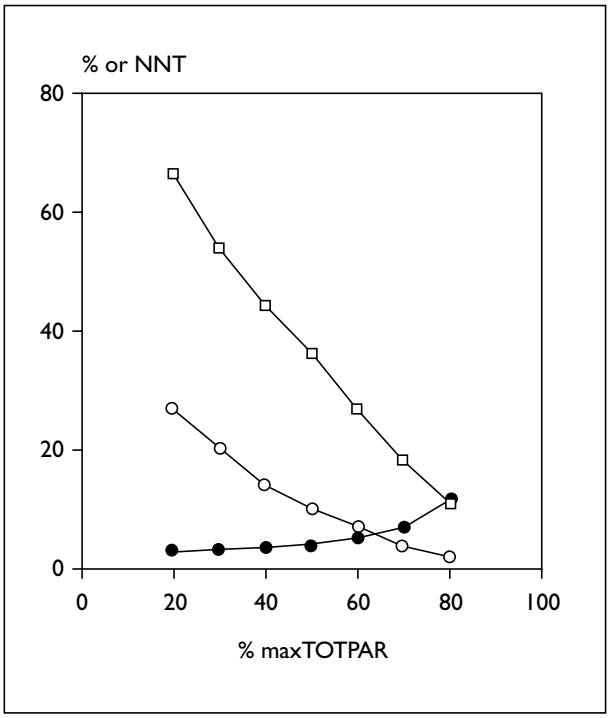


FIGURE 22 *Effect of different thresholds of pain relief on NNT (–□–, paracetamol plus propoxyphene; –●–, placebo; –○–, NNT)*

target was raised. For an effective analgesic, this proportion fell slowly until high relief targets were reached. The difference remained largely unaltered over a wide range of targets, thus generating stable NNTs.

An NNT of 1 describes an event which occurs in every patient given the treatment but in no patient in a comparator group. This could be described as the 'perfect' result in, say, a therapeutic trial of an antibiotic compared with placebo. For therapeutic benefit, the NNT value should be as close as possible to 1; there are few circumstances in which a treatment is close to 100% effective and the control or placebo completely ineffective, so NNTs of 2 or 3 often indicate an effective intervention. For unwanted effects, NNT becomes the NNH (number-needed-to-harm), which should be as large as possible.

It is important to remember that the NNT is always relative to the comparator and applies to a particular clinical outcome. The duration of treatment necessary to achieve the target should be specified. The NNT for cure of head-lice at 2 weeks with permethrin 1% compared with a control was 1.1 (95% CI, 1.0–1.2).[2,19]

Confidence intervals

The CI of the NNT is an indication that, in 19 cases out of 20, the 'true' value will be in the specified range. If the odds ratio or relative risk/benefit is not statistically significant then the NNT is infinite, indicating no difference from control. An NNT with an infinite CI is then but a point estimate. It may still have clinical importance as a benchmark until further data permits finite CIs but decisions must take account of this parlous state.

Disadvantages

The disadvantage of the NNT approach, apparent from the formula, is that it needs dichotomous data. Continuous data can be converted to dichotomous for acute pain studies so that NNTs may be calculated by deriving a relationship between the two from individual patient data.[20] Because of the way in which it is calculated, NNT will also be sensitive to trials with high CERs. As the CER rises, the potential for treatment-specific improvement decreases, resulting in higher (and apparently less effective) NNTs. So, as with any summary measure from a quantitative systematic review, NNT needs to be treated with caution; comparisons can only be made confidently if CERs are in the same range.

Calculating NNTs when they are not provided

Odds ratios

If a quantitative systematic review produces odds ratios but no NNTs, these can be derived from *Table 18*.

A caveat must be added here that odds ratios should be interpreted with caution when events occur commonly, such as in treatments, and odds ratios may overestimate the benefits of an effect when event rates are above 50%. They are likely to be superseded by relative risk or benefit because these are more robust in situations where event rates are high.[14,21]

Is it safe?

Estimating the risk of harm is a critical part of a clinical decision. Systematic reviews should report adverse events as well as efficacy, and consider the issue of rare but important adverse events. Large RCTs apart, most trials study limited numbers of patients. New medicines may be launched after trials on 1500 patients,[22] missing any rare but important adverse events. The rule of three is important here. If a particular serious event does not occur in 1500 patients given the treatment, then we can be 95% confident that the chance of it occurring is, at most, 3/1500.[23]

TABLE 18 *Table for estimating NNT when odds ratio or CER are known (for prophylactic interventions[16])*

	Odds ratios																		
	Prophylaxis									Treatment									
CER	**0.5**	**0.55**	**0.6**	**0.65**	**0.7**	**0.75**	**0.8**	**0.85**	**0.9**	**1.5**	**2.0**	**2.5**	**3.0**	**3.5**	**4.0**	**4.5**	**5.0**	**10.0**
0.05	41	46	52	59	69	83	104	139	209	43	22	15	12	9	8	7	6	3
0.1	21	24	27	31	36	43	54	73	110	23	12	9	7	6	5	4	4	2
0.2	11	13	14	17	20	24	30	40	61	14	8	5	4	4	3	3	3	2
0.3	8	9	10	12	14	18	22	30	46	11	6	5	4	3	3	3	3	2
0.4	7	8	9	10	12	15	19	26	40	10	6	4	4	3	3	3	3	2
0.5	6	7	8	9	11	14	18	25	38	10	6	5	4	4	3	3	3	2
0.7	6	7	9	10	13	16	20	28	44	13	8	7	6	5	5	5	5	4
0.9	12	15	18	22	27	34	46	64	101	32	21	17	16	14	14	13	13	11

Formula for **prophylaxis**:

$$\frac{\text{NNT} = 1 - [\text{CER} \times (1 - \text{OR})]}{(1 - \text{CER}) \times \text{CER} \times (1 - \text{OR})}$$

Formula for **treatment**:

$$\frac{\text{NNT} = \text{CER}\ (\text{OR} - 1) + 1}{\text{CER}\ (\text{OR} - 1) \times (1 - \text{CER})}$$

where OR = odds ratio.

Choose the column which is closest to the published odds ratio (**prophylaxis** left side, **treatment** right side) and the row which is closest to the expected CER, then read off the corresponding NNT. The table can also be used to see how different values for CER or EER for an individual patient affect the NNT at a given odds ratio.

Much the same rules apply to harm as to efficacy, but with some important differences – the NNH rather than the NNT and the rules of admissible evidence.

Number-needed-to-harm

For minor adverse effects reported in RCTs, NNH may be calculated in the same way as the NNT. When there is low incidence it is likely that point estimates alone will emerge (infinite CIs). Major harm may be defined in a set of RCTs as intervention-related study withdrawal, and be calculated from those numbers. Precise estimates of major harm will require a much wider literature search to trawl for case reports or series. The absence of information on adverse effects in systematic reviews reduces their usefulness.

Rules of admissible evidence

The gold standard of evidence for harm, as for efficacy, is the RCT. The problem is that, in the relatively small number of patients studied in RCTs, rare serious harm may not be spotted. For an adverse effect systematic review, study architectures of lower intrinsic quality may therefore be admissible. An extreme example is that observer blinding is superfluous if the outcome is death. Such rare and serious harm cannot and should not be dismissed just because it is reported in a case report rather than in an RCT. The 'process rules' in this area have yet to be determined.

Using NNTs

In an ideal world you will have three numbers for each intervention, an NNT for benefit and an NNH for both minor and major harm.

This then becomes the yardstick against which alternative interventions should be judged, and the pivot for the clinical decision on whether or not to use the intervention for an individual patient.

References

1. Oxman AD, Guyatt GH. Guidelines for reading literature reviews. *Can Med Assoc J* 1988;**138**:697–703.
2. Vander Stichele RH, Dezeure EM, Bogaert MG. Systematic review of clinical efficacy of topical treatments for head lice. *BMJ* 1995;**311**:604–8.
3. Schulz KF, Chalmers I, Hayes RJ, Altman DG. Empirical evidence of bias: dimensions of methodological quality associated with estimates of treatment effects in controlled trials. *JAMA* 1995;**273**:408–12.
4. Carroll D, Tramèr M, McQuay H, Nye B, Moore A. Randomization is important in studies with pain outcomes: systematic review of transcutaneous electrical nerve stimulation in acute postoperative pain. *Br J Anaesth* 1996;**77**:798–803.
5. Jadad AR, Moore RA, Carroll D, Jenkinson C, Reynolds DJM, Gavaghan DJ, *et al.* Assessing the quality of reports of randomized clinical trials: is blinding necessary? *Control Clin Trials* 1996;**17**:1–12.
6. L'Abbé KA, Detsky AS, O'Rourke K. Meta-analysis in clinical research. *Ann Intern Med* 1987;**107**:224–33.
7. Zhang WY, Li WPA. The effectiveness of topically applied capsaicin. A meta-analysis. *Eur J Clin Pharmacol* 1994;**46**:517–22.
8. McQuay H, Carroll D, Jadad AR, Wiffen P, Moore A. Anticonvulsant drugs for management of pain: a systematic review. *BMJ* 1995;**311**:1047–52.
9. McQuay HJ, Tramèr M, Nye BA, Carroll D, Wiffen PJ, Moore RA. A systematic review of antidepressants in neuropathic pain. *Pain* 1996;**68**:217–27.
10. Tramèr M, Moore A, McQuay H. Prevention of vomiting after paediatric strabismus surgery: a systematic review using the numbers-needed-to-treat method. *Br J Anaesth* 1995;**75**:556–61.
11. Soll JC, McQueen MC. Respiratory distress syndrome. In: Sinclair JC, Bracken ME, editors. Effective care of the newborn infant. Oxford: Oxford University Press, 1992:chapter 15, p333.
12. Moore RA. *Helicobacter pylori* and peptic ulcer. A systematic review of effectiveness and an overview of the economic benefits of implementing that which is known to be effective (http://www.jr2.ox.ac.uk/Bandolier/bandopubs/hpyl/hp0.html). Oxford: Health Technology Evaluation Association, 1995.
13. Gøtzsche PC. Sensitivity of effect variables in rheumatoid arthritis: a meta analysis of 130 placebo controlled NSAID trials. *J Clin Epidemiol* 1990;**43**:1313–18.
14. Sinclair JC, Bracken MB. Clinically useful measures of effect in binary analyses of randomized trials. *J Clin Epidemiol* 1994;**47**:881–9.
15. Tramèr M, Moore RA, McQuay HJ. A rational approach to sensitivity analyses in meta-analysis: factors influencing propofol's effect on postoperative nausea and vomiting. Submitted 1997.
16. Sackett DL, Deeks JJ, Altman DG. Down with odds ratios! *Evidence-Based Med* 1996;**1**:164–6.
17. Deeks J. What the heck's an odds ratio? *Bandolier* 1996;**3**(25):6–7.
18. Glass GV. Primary, secondary, and meta-analysis of research. *Educ Res* 1976;**5**:3–8.
19. Moore RA, McQuay H, Gray M. Bandolier – the first 20 issues. Oxford: Bandolier, 1995.
20. Moore A, McQuay H, Gavaghan D. Deriving dichotomous outcome measures from continuous data in randomised controlled trials of analgesics. *Pain* 1996;**66**:229–37.
21. Sackett D, Richardson WS, Rosenberg W, Haynes B. Evidence based medicine. London: Churchill Livingstone, 1996.
22. Moore TJ. Deadly medicine. New York: Simon & Schuster, 1995.
23. Eypasch E, Lefering R, Kum CK, Troidl H. Probability of adverse events that have not yet occurred: a statistical reminder. *BMJ* 1995;**311**:619–20.

Chapter 7

Existing systematic reviews

As part of the evidence-gathering exercise our aim was to find all previous systematic reviews of analgesic interventions.

This was undertaken in two stages. In the first, a search was undertaken for all systematic reviews published before 1993/94. A total of 80 were found, and two judges assessed their quality using Oxman and Guyatt's index.[1] Most of the reviews looked at drug interventions for chronic pain conditions. Two-thirds were published after 1990. Most had methodological flaws, such as insufficient information on retrieval methods and validity assessment and design of the primary studies. Poor quality systematic reviews reached significantly more positive conclusions and, when there was more than one systematic review on a particular topic, the results did not always agree. A full account of this first stage has been published.[2]

The second stage was (and is) a prospective exercise to maintain an up-to-date database of systematic reviews in pain relief.

Introduction

Systematic reviews can potentially resolve conflicts when reports of primary studies disagree, and increase the likelihood of detecting small but clinically important effects.[3–5] They can also be easily misused to produce misleading estimates of effectiveness.

A systematic search of the literature was used to identify the highest possible proportion of systematic reviews assessing analgesic interventions. The objectives were:

(i) to produce a citation database of all available reviews
(ii) to assess the quality of systematic reviews in pain relief
(iii) to establish whether or not quality scores are useful to resolve conflicts between different systematic reviews.

Methods

Inclusion criteria

Reports had to meet the following criteria.

1. They had to be described as systematic reviews or, if not, they had to include pooled analysis of the results of several independent primary studies. Studies in which statistical synthesis had been planned but was deemed to be inappropriate were also included.
2. They had to incorporate trials in which pain was an outcome measure or in which analgesic interventions were compared for outcomes other than pain within the context of a painful condition (e.g. a study looking at the validity of grip strength to assess the effectiveness of NSAIDs in rheumatoid arthritis).
3. They had to be published or accepted for publication.

Search strategy

A MEDLINE search (Silver Platter MEDLINE v. 3.0, 3.1 and 3.11) was undertaken from 1966 to October 1993. This MEDLINE strategy had been developed to identify the maximum possible number of randomized, double-blind studies or meta-analyses in pain research, and contained text words, 'wild cards' and MeSH terms.[6] Forty journals were searched by hand. The register of systematic reviews at the UK Cochrane Centre was checked for eligible studies, and lead authors of abstracts were asked for full manuscripts. The reference lists for citations of other systematic reviews were scanned.

Methodological evaluation

Each study was evaluated twice, using Oxman and Guyatt's index,[1,7] with the title of the journal, the authors' names, the date of publication and the source of financial support for the study obscured. A consensus score was obtained.

Statistical analysis

The chi-squared test was used to test the relationship between the direction of the conclusion of the systematic reviews (positive versus negative/uncertain) and the overall quality scores, and the influence of study architecture on systematic reviews which included study designs other than RCTs. Prior hypotheses were that poor quality reviews and those including designs other than RCTs would be more likely to produce positive conclusions.

Results

Quality assessment of reviews to 1993/94

Of the 84 reports found, 70 were included in the quality assessment *(Table 19)*. The exclusions are specified elsewhere.[2] The earliest report was from 1980, and over two-thirds appeared after 1990. Reviews considered between two and 196 primary studies (median 28). In all, 60 reviews reached positive conclusions, seven negative conclusions, 12 were uncertain and one did not reach any conclusion. They used different pooling methods *(Table 20)*. All were based on published data only (no individual patient data analysis), without validity checks with the study investigators. A separate list of the reviews, by first author, is presented at the end of this chapter.

Overall quality scores

The median agreed overall score for the systematic reviews was 4 (range 1–7). Systematic reviews of high quality were significantly less likely to produce positive results (see *Table 21* and *Figure 23;* chi-squared 18.2, $p = 0.006$).

TABLE 19 *Details of the systematic reviews*

Setting	Number (%)	Outcomes	Number (%)
Chronic	58 (72)	Pain	63 (79)
Acute	14 (19)	Adverse effects	11 (14)
Mixed	6 (7)	Validity	3 (4)
Unclear	2 (2)	Patient preference	1 (1)
Intervention		Return to work	1 (1)
Drug	42 (54)	Pulmonary function	1 (1)
Psychological	16 (20)	**Primary studies**	
Physical	10 (13)	Randomised only	24 (30)
Diagnostic	3 (4)	Randomised and double-blind	7 (9)
Complementary	2 (2)	Double-blind only	3 (4)
Non-surgical invasive	2 (2)	Combination of observational and any of the above	25 (31)
Multidisciplinary	2 (2)	Observational only	4 (5)
Surgical	1 (1)	Not reported	17 (21)
Preventive	1 (1)		
Not specified	1 (1)		

TABLE 20 *Pooling methods used in the systematic reviews*

Method	Number (%)	Method	Number (%)
Standardised mean differences	26 (32)	Random effects	2 (2)
Odds ratios/Mantel–Haenszel	10 (13)	Kendal's correlation	1 (1)
'Percentage change' comparison	15 (20)	Log rank test	1 (1)
Simple addition	7 (9)	Relative potency	1 (1)
Criteria-based	4 (5)	Not reported	5 (6)
Weighted means	3 (4)	Pooling considered inappropriate	3 (4)
Mean risk differences	2 (2)		

TABLE 21 *Meta-analyses: quality and conclusions*

Overall quality score	Conclusion	
	Positive	Negative/ Uncertain
1	11	1
2	12	0
3	6	1
4	11	1
5	8	3
6	8	9
7	4	4

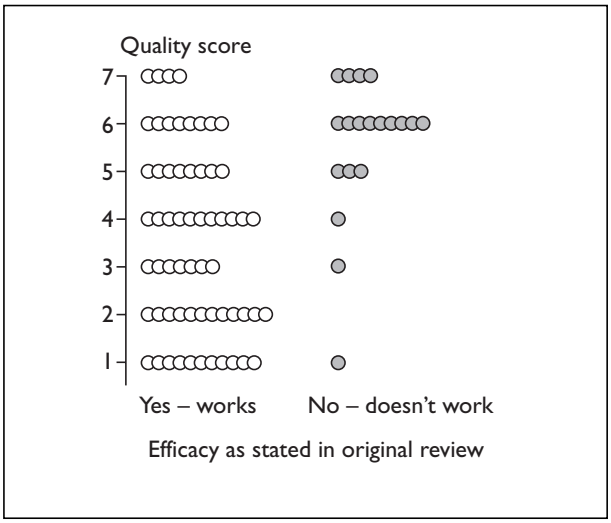


FIGURE 23 *Systematic reviews: quality and estimate of efficacy*[2]

Of 19 systematic reviews with negative or uncertain results, 16 had overall quality scores above the median, compared with only 20 of the 60 with positive results. Systematic reviews restricted to RCTs were significantly less likely to produce positive conclusions (19/31) than those which included other study architectures (41/49; chi-squared = 5.07; $p = 0.024$). All conclusions from systematic reviews of psychological interventions were positive. In only one of those reviews was quality scored above the median. All abstracts scored below the median, and six out of eight abstracts received the minimum possible score.

Interventions evaluated by multiple systematic reviews

There was more than one systematic review for six interventions *(Table 22)*. For acupuncture and NSAIDs the conclusions of the reviews were the same. Two reviews of acupuncture in chronic pain concluded that the evidence was flawed and that acupuncture was of uncertain value.[8,9] Three reviews confirmed that the risk of gastrointestinal complications was increased by NSAIDs.[10–12]

Most systematic reviews of manipulation for chronic back pain concluded that it was useful,[13–15] as did

TABLE 22 *Multiple systematic reviews on a particular intervention: quality and conclusions*

Intervention	Study	Quality score	Conclusion
Acupuncture	ter Riet, *et al.*[9]	6	Uncertain
	Patel, *et al.*[8]	5	Uncertain
Gastrointestinal effects of NSAIDs	Chalmers, *et al.*[10]	7	Positive
	Gabriel, *et al.*[11]	6	Positive
	Bollini, *et al.*[12]	5	Positive
Manipulation	Koes, *et al.*[19]	6	Uncertain
	Shekelle, *et al.*[13]	6	Positive
	Ottenbacher & di Fabio[14]	4	Positive
	Anderson, *et al.*[15]	3	Positive
Second-line drugs for rheumatoid arthritis	Gøtzsche, *et al.*[20]	7	Uncertain
	Felson, *et al.*[17]	6	Positive
	Capell, *et al.*[16]	4	Positive
	Felson, *et al.*[18]	4	Positive
Prevention of postherpetic neuralgia	Schmader & Studenski[23]	7	Uncertain
	Lycka[24]	5	Positive
	Crooks, *et al.*[26]	3	Positive
	Naldi, *et al.*[25]	3	Negative
Laser for musculoskeletal pain	Gam, *et al.*[22]	6	Negative
	Beckerman, *et al.*[21]	6	Positive

BOX 2 Using the Oxman and Guyatt scoring system for reviews

The purpose of this index is to evaluate the scientific quality (i.e. adherence to scientific principles) of research overviews (review articles) published in the medical literature. It is not intended to measure literary quality, importance, relevance, originality, or other attributes of overviews.

The index is for assessing overviews of primary ('original') research on pragmatic questions regarding causation, diagnosis, prognosis, therapy, or prevention. A research overview is a survey of research. The same principles that apply to epidemiological surveys apply to overviews; a question must be clearly specified, a target population identified and accessed, appropriate information obtained from that population in an unbiased fashion, and conclusions derived, sometimes with the help of formal statistical analysis, as in 'meta-analyses'. The fundamental difference between overviews and epidemiological surveys is the unit of analysis, not the scientific issues addressed by the questions in this index.

Since most published overviews do not include a methods section, it is difficult to answer some of the questions in the index. Answers need to be based, as far as possible, on information provided in the overview. If the methods used are incompletely reported relative to a specific item, that item should be scored as 'partially'. Similarly, if no information is provided regarding what was done relative to a particular question, it should be scored as 'can't tell', unless there is information in the overview to suggest that the criterion either was or was not met.

Quality features

1	Were the search methods used to find evidence on the primary question(s) stated?	No	Partially	Yes
2	Was the search for evidence reasonably comprehensive?	No	Can't tell	Yes
3	Were the criteria used for deciding which studies to include in the overview reported?	No	Partially	Yes
4	Was bias in the selection of studies avoided?	No	Can't tell	Yes
5	Were the criteria used for assessing the validity of the included studies reported?	No	Partially	Yes
6	Was the validity of all studies referred to in the text assessed using appropriate criteria?	No	Can't tell	Yes
7	Were the methods used to combine the findings of the relevant studies (to reach a conclusion) reported?	No	Partially	Yes
8	Were the findings of the relevant studies combined appropriately relative to the primary question of the overview?	No	Can't tell	Yes
9	Were the conclusions reached by the author(s) supported by the data and/or analysis reported in the overview?	No	Partially	Yes
10	How would you rate the scientific quality of this overview?			

Flaws

Extensive		Major		Minor	Minimal	
1	2	3	4	5	6	7

For Question 8, if no attempt has been made to combine findings, and no statement is made regarding the inappropriateness of combining findings, check 'no'. If a summary (general) estimate is given anywhere in the abstract, the discussion, or the summary section of the paper, and it is not reported how that estimate was derived, check 'no', even if there is a statement regarding the limitations of combining the findings of the studies reviewed. If in doubt mark as 'can't tell'.

For Question 9, if an overview is to be scored as 'yes', data (not just citations) must be reported that support the main conclusions relating to the primary question(s) addressed by the overview.

For Question 10, overall scientific quality, the score should be based on the answers to the first nine questions. The following guidelines can be used to assist with deriving a summary score: If the 'can't tell' option is used one or more times on the preceding questions, a review is likely to have minor flaws at best and it is difficult to rule out major flaws (i.e. a score of 4 or less). If Questions 2, 4, 6 or 8 are marked as 'no', the review is likely to have major flaws (i.e. a score of 3 or less, depending on the number and degree of the flaws).

reviews of second-line drugs for rheumatoid arthritis,[16–18] but for both interventions one review questioned the validity of the findings because of the high risk of bias in the primary studies.[19,20]

Systematic reviews produced conclusions in opposite directions for lasers in musculoskeletal pain[21,22] and for interventions to prevent post-herpetic neuralgia.[23–26] Both systematic reviews evaluating laser treatment were given the same quality score.

Comments

The use of systematic reviews to assess analgesic interventions is increasing but most of the reviews found had methodological flaws which may threaten their conclusions. Only eight of the 80 satisfied all the Oxman and Guyatt criteria (see *Box 2* for explanation) and 16% were given the lowest possible score. The relationship between methodological rigour, type of primary studies included and the direction of the conclusions underscores the importance of review quality. Systematic reviews including only RCTs were less likely to produce positive conclusions.

Reviewers have to work hard to reduce bias. The search for evidence must be comprehensive, decisions about which studies to include or exclude have to be overt, and validity criteria need to be stated. Equally, readers need to be aware of the pitfalls.

Several examples were found of reviews of the same intervention producing conflicting results, despite similar quality scores. This despite the concept that systematic reviews can resolve conflicting results between primary studies.

References

1. Oxman AD, Guyatt GH. Validation of an index of the quality of review articles. *J Clin Epidemiol* 1991;**44**:1271–8.
2. Jadad AR, McQuay HJ. Meta-analyses to evaluate analgesic interventions: a systematic qualitative review of their methodology. *J Clin Epidemiol* 1996;**49**:235–43.
3. Sacks HS, Berrier J, Reitman D, Ancona-Berk VA, Chalmers TC. Meta-analyses of randomized controlled trials. *N Engl J Med* 1987;**316**:450–5.
4. Gerbarg ZB, Horwitz RI. Resolving conflicting clinical trials: guidelines for meta-analysis. *J Clin Epidemiol* 1988;**41**:503–9.
5. Kassirer JP. Clinical trials and meta-analysis. What do they do for us? *N Engl J Med* 1992;**327**:273–4.
6. Jadad AR, McQuay HJ. A high-yield strategy to identify randomized controlled trials for systematic reviews. *Online J Curr Clin Trials* [serial online] 1993; Doc No 33 [3973 words; 39 paragraphs; 5 tables].
7. Oxman AD, Guyatt GH, Singer J, Goldsmith CH, Hutchison BG, Milner RA, *et al.* Agreement among reviewers of review articles. *J Clin Epidemiol* 1991;**44**:91–8.
8. Patel M, Gutzwiller F, Paccaud F, Marazzi A. A meta-analysis of acupuncture for chronic pain. *Int J Epidemiol* 1989;**18**:900–6.
9. ter Riet G, Kleijnen J, Knipschild P. Acupuncture and chronic pain: a criteria-based meta-analysis. *J Clin Epidemiol* 1990;**43**:1191–9.
10. Chalmers TC, Berrier J, Hewitt P, Berlin J, Reitman D, Nagalingam R, *et al.* Meta-analysis of randomized controlled trials as a method of estimating rare complications of non-steroidal anti-inflammatory drug therapy. *Aliment Pharmacol Therap* 1988;**2**:9–26.
11. Gabriel SE, Jaakkimainen L, Bombardier C. Risk for serious gastrointestinal complications related to use of nonsteroidal anti-inflammatory drugs. A meta-analysis. *Ann Intern Med* 1991;**115**:787–96.
12. Bollini P, García Rodríguez LA, Pérez Gutthann S, Walker AM. The impact of research quality and study design on epidemiologic estimates of the effect of nonsteroidal anti-inflammatory drugs on upper gastrointestinal tract disease. *Arch Intern Med* 1992;**152**:1289–95.
13. Shekelle PG, Adams AH, Chassin MR, Hurwitz EL, Brook RH. Spinal manipulation for low-back pain. *Ann Intern Med* 1992;**117**:590–8.
14. Ottenbacher K, di Fabio RP. Efficacy of spinal manipulation/mobilization therapy. A meta analysis. *Spine* 1985;**10**:833–7.
15. Anderson R, Meeker WC, Wirick BE, Mootz RD, Kirk DH, Adams A. A meta-analysis of clinical trials of spinal manipulation. *J Manipulative Physiol Ther* 1992;**15**:181–94.
16. Capell HA, Porter DR, Madhok R, Hunter JA. Second line (disease modifying) treatment in rheumatoid arthritis: which drug for which patient? *Ann Rheum Dis* 1993;**52**:423–8.
17. Felson DT, Anderson JJ, Meenan RF. The comparative efficacy and toxicity of second-line drugs in rheumatoid arthritis. *Arthritis Rheum* 1990;**33**:1449–61.
18. Felson DT, Anderson JJ, Meenan RF. Use of short-term efficacy/toxicity trade offs to select second-line drugs in rheumatoid arthritis. *Arthritis Rheum* 1992;**35**:1117–25.
19. Koes BW, Assendelft WJ, van der Heijden GJ, Bouter LM, Knipschild PG. Spinal manipulation and mobilisation for back and neck pain: a blinded review. *BMJ* 1991;**303**:1298–303.

20. Gøtzsche PC, Pødenphant J, Olesen M, Halberg P. Meta-analysis of second-line antirheumatic drugs: sample bias and uncertain benefit. *J Clin Epidemiol* 1992;**45**:587–94.

21. Beckerman H, de Bie RA, Bouter LM, de Cuyper HJ, Oostendorp RA. The efficacy of laser therapy for musculoskeletal and skin disorders: a criteria-based meta-analysis of randomized clinical trials. *Phys Ther* 1992;**72**:483–91.

22. Gam AN, Thorsen H, Lonnberg F. The effect of low-level laser therapy on musculoskeletal pain: a meta-analysis. *Pain* 1993;**52**:63–6.

23. Schmader K, Studenski S. Are current therapies useful for the prevention of postherpetic neuralgia? A critical analysis of the literature. *J Gen Intern Med* 1989;**4**:83–9.

24. Lycka BA. Postherpetic neuralgia and systemic corticosteroid therapy. Efficacy and safety. *Int J Dermatol* 1990;**29**:523–7.

25. Naldi L, Zucchi A, Brevi A, Cavalieri d'Oro L, Cainelli T. Corticosteroids and post-herpetic neuralgia. *Lancet* 1990;**336**:947.

26. Crooks RJ, Jones Da, Fiddian P. Zoster-associated chronic pain: an overview of clinical trials with acyclovir. *Scand J Infect* 1991;**78** suppl:62–8.

Reviews in chronic pain

Abenhaim L, Bergeron AM, 1992. Twenty years of randomized clinical trials of manipulative therapy for back pain: a review. *Clin Invest Med*;**15**:527–35.

Abram SE, 1993. Advances in chronic pain management since gate control [1992 Bonica lecture]. *Reg Anesth*;**18**:66–81.

Adamson GD, Nelson HP, 1996. Medical and surgical treatment of endometriosis. *Endocrinologist*;**6**:384–91.

Aker PD, Gross AR, Goldsmith CH, Peloso P, 1996. Conservative management of mechanical neck pain: systematic overview and meta-analysis. *BMJ*;**313**:1291–6.

Ali NMK, 1995. Does sympathetic ganglionic block prevent postherpetic neuralgia? *Reg Anesth*;**20**:227–33.

Allegrante JP, 1996. The role of adjunctive therapy in the management of chronic nonmalignant pain. *Am J Med;***101**:33–39S.

Altman GB, Lee CA, 1996. Strontium-89 for treatment of painful bone metastasis from prostate cancer. *Oncol Nurs Forum*;**23**:523–7.

Amery WK, 1988. Onset of action of various migraine prophylactics. *Cephalalgia*;**8**:11–13.

Amery WK, 1989. Prophylactic and curative treatment of migraine with calcium antagonists. *Drug Des Deliv*;**4**:197–203.

Anderson R, Meeker WC, Wirick BE, Mootz RD, Kirk DH, Adams A, 1992. A meta-analysis of clinical trials of spinal manipulation. *J Manipulative Physiol Ther*;**15**:181–94.

Antczak-Bouckoms A, Tung F, Chalmers TC, Bouckoms A, 1987. Quality assessment and meta-analysis of randomized controlled trials of ibuprofen. *Pain*;suppl:S51.

Assendelft WJ, Koes BW, van der Heijden GJ, Bouter LM, 1992. The efficacy of chiropractic manipulation for back pain: blinded review of relevant randomized clinical trials. *J Manipulative Physiol Ther*;**15**:487–94.

Assendelft WJJ, Hay EM, Adshead R, Bouter LM, 1996. Corticosteroid injections for lateral epicondylitis: a systematic overview. *Br J Gen Pract;***46**:209–16.

Assendelft WJJ, Koes BW, van der Heijden GJM, Bouter LM, 1996. The effectiveness of chiropractic treatment of low back pain: an update and attempt at statistical pooling. *J Manipulative Physiol Ther*;**19**:499–507.

Aylward M, 1975. Clinical studies on alclofenac in the treatment of rheumatic diseases: a drug in question. *Curr Med Res Opin*;**3**:274–85.

Ballantyne JC, Carr DB, Chalmers TC, Dear KB, Angelillo IF, Mosteller F, 1993. Postoperative patient-controlled analgesia: meta-analyses of initial randomized control trials. *J Clin Anesth*;**5**:182–93.

Ballantyne JC, Carr DB, Berkey CS, Chalmers TC, Mosteller F, 1996. Comparative efficacy of epidural, subarachnoid and intracerebroventricular opioids in patients with pain due to cance. *Reg Anesth*;**21**:542–56.

Bamberg P, Caswell CM., Frame MH, Lam SK, Wong EC, 1992. A meta-analysis comparing the efficacy of omeprazole with H_2-receptor antagonists for acute treatment of duodenal ulcer in Asian patients. *J Gastroenterol Hepatol*;**7**:577–85.

Baron JC, Bousser MG, 1982. Essais therapeutiques controles dans la migraine. Organisation et resultats [Controlled therapeutic trials in migraine: methodology and results]. *Rev Neurol Paris*;**138**:279–95.

Basinski A, Naylor CD, 1991. Aspirin and fibrinolysis in acute myocardial infarction: meta-analytic evidence for synergy. *J Clin Epidemiol*;**44**:1085–96.

Beckerman H, de Bie RA, Bouter LM, De Cuyper HJ, Oostendorp RA, 1992. The efficacy of laser therapy for musculoskeletal and skin disorders: a criteria-based meta-analysis of randomized clinical trials. *Phys Ther*;**72**:483–91.

Beecher HK, 1955. The powerful placebo. *JAMA*;**159**:1602–6.

Belcon MC, Rooney PJ, Tugwell P, 1985. Aspirin and gastrointestinal haemorrhage: a methodologic assessment. *J Chronic Dis*;**38**:101–11.

Bernstein J, 1983. Overview of efficacy of fenbufen in rheumatoid arthritis and osteoarthritis. *Am J Med*;**75**:70–4.

Bertagna C, de Gery A, Hucher M, Francois JP, Zanirato J, 1994. Efficacy of the combination of nilutamide plus orchidectomy in patients with metastatic prostatic cancer. A meta-analysis of seven randomized double-blind trials (1056 patients). *Br J Urol*;**73**:396–402.

Bhatt Sanders D, 1985. Acupuncture for rheumatoid arthritis: an analysis of the literature. *Semin Arthritis Rheum*;**14**:225–31.

Blanchard EB, Andrasik F, 1982. Psychological assessment and treatment of headache: recent developments and emerging issues. *J Consult Clin Psychol*;**50**:859–79.

Blanchard EB, Andrasik F, Ahles TA, Teders ST, O'Keefe D, 1980. Migraine and tension headache: a meta-analytic review. *Behav Ther*;**11**:613–31.

Bloomfield SS, Mitchell J, Cissell G, Barden TP, 1986. Analgesic sensitivity of two postpartum pain models. *Pain*;**27**:171–9.

Bogaards MC, ter Kuile MM, 1994. Treatment of recurrent tension headache: a meta-analytic review. *Clin J Pain*;**10**:174–90.

Bollini P, García Rodríguez LA, Pérez Gutthann S, Walker AM. The impact of research quality and study design on epidemiologic estimates of the effect of non-steroidal anti-inflammatory drugs on upper gastrointestinal tract disease. *Arch Intern Med*;**152**:1289–95.

Bosi Ferraz M, Atra E, 1989. Meta-análise: nova opcao em pesquisa clínica. Prinípos básicos e um exemplo clínico. *Rev Paul Med*;**107**:5–9.

Bosi-Ferraz M, Tugwell P, Goldsmith CH, Atra E, 1990. Meta-analysis of sulfasalazine in ankylosing spondylitis. *J Rheumatol*;**17**:1482–6.

Broome ME, Lillis PP, 1989. A descriptive analysis of the pediatric pain management research. *Appl Nurs Res*;**2**:74–81.

Broome ME, Lillis PP, Smith MC, 1989. Pain interventions with children: a meta analysis of research. *Nurs Res*;**38**:154–8.

Brown BWJ, 1993. Meta-analysis and patient-controlled analgesia [editorial; comment]. *J Clin Anesth*;**5**:179–81.

Brunarski DJ, 1984. Clinical trials of spinal manipulation: a critical appraisal and review of the literature. *J Manipulative Physiol Ther*;**7**:243–9.

Capell HA, Porter DR, Madhok R, Hunter JA, 1993. Second line (disease modifying) treatment in rheumatoid arthritis: which drug for which patient? *Ann Rheum Dis*;**52**:423–8.

Capurso L, Koch M, 1991. La prevenzione del danno gastrico da FANS: H2 antagonisti o misoprostol? Una meta-analisi degli studi clinic controllati [Prevention of NSAID-induced gastric lesions: H2 antagonists or misoprostol? A meta-analysis of controlled clinical studies]. *Clin Ter*;**139**:179–89.

Carr D, Eisenberg E, Chalmers TC, 1993. Neurolytic celiac plexus block for cancer pain – a meta-analysis [abstract]. In: Proceedings 7th World Congress on Pain. Oxford: IASP/Elsevier;338.

Carroll D, Tramèr M, McQuay H, Nye B, Moore A, 1996 Randomization is important in studies with pain outcomes: systematic review of trancutaneous electrical nerve stimulation in acute postoperative pain. *Br J Anaesth*;**77**:798–803.

Carroll D, Tramèr M, McQuay H, Nye B, Moore A, 1997. Transcutaneous electrical nerve stimulation in labour pain: a systematic review. *Br J Obstet Gynaecol*;**104**:169–75.

Carroll D, Moore RA, Tramèr MR, McQuay HJ, 1998. Transcutaneous electrical nerve stimulation does not relieve labour pain: updated systematic review. *Contemp Rev Obstet Gynaecol* 1997;(Sept):195–205.

Chalmers TC, Berrier J, Bewitt P, Berlin J, Reitman D, Nagalingam R, *et al.*, 1988. Meta-analysis of randomized controlled trials as a method of estimating rare complications of non-steroidal anti-inflammatory drug therapy. *Aliment Pharmacol Ther*;**2**:9–26.

Chalmers TC, Ballantyne JC, Carr DB, Dear KBG, Angelillo IF, Mosteller F, 1993. Comparative effects of analgesic therapies upon postoperative pulmonary function: meta-analyses [abstract]. In: Proceeding 7th World Congress on Pain. Oxford: IASP/Elsevier;136.

Christensen E, Juhl E, Tygstrup N, 1977. Treatment of duodenal ulcer. Randomized clinical trials of a decade (1964 to 1974). *Gastroenterology*;**73**:1170–8.

Cohen JE, Goel V, Frank JW, Bombardier C, Peloso P, Guillemin F, 1994. Group education interventions for people with low back pain. An overview of the literature. *Spine*;**19**:1214–22.

Collins S, Moore A, McQuay H, 1996. Paracetamol-codeine combinations versus paracetamol alone. Actual size of increase needs to be measured [letter]. *BMJ*;**313**:1209.

Collins SL, Moore A, McQuay HJ, Wiffen PJ, 1999. Oral ibuprofen and diclofenac in postoperative pain: a quantitative systematic review. *Eur J Pain*. In press.

Collins SL, Edwards J, Moore A, McQuay HJ, 1998. Single dose oral dextropropoxyphene in postoperative pain: a quantitative systematic review. *Eur J Clin Pharmacol*;**54**:107–12.

Cooper SA, 1984. Five studies on ibuprofen for postsurgical dental pain. *Am J Med*;suppl:70–7.

Cooper SA, Needle SE, Kruger GO, 1977. Comparative analgesic potency of aspirin and ibuprofen. *J Oral Surg*;**35**:898–903.

Cooper SA, Gross DJ, Vogel R, Berrie P, 1983. Relative analgesic efficacy of ibuprofen and zomepirac Na in periodontal and oral surgery. *Clin Pharmacol Ther*;**33**:194.

Cooper SA, Precheur H, Rosenbeck R, Engel J, 1984. The analgesic efficacy of ibuprofen compared to acetaminophen with codeine. *Clin Pharmacol Ther*;**35**:232.

Cooper SA, Berrie R, Cohn P, 1988. Comparison of ketoprofen, ibuprofen, and placebo in a dental surgery pain model. *Adv Ther*;**5**:43–53.

Crooks RJ, Jones Da, Fiddian P, 1991. Zoster-associated chronic pain: an overview of clinical trials with acyclovir. *Scand J Infect*;**78** suppl:62–8.

Cutler R, Fishbain D, Rosomoff H, Abdel-Moty E, Khalil T, Steele-Rosomoff R, 1992. Does non-surgical pain treatment of chronic pain return patients to work? In: Proceedings of American Pain Society Meeting 1992:83.

Cutler RB, Fishbain DA, Rosomoff HL, Abdel-Moty E, Khalil TM, Rosomoff RS, 1994. Does nonsurgical pain center treatment of chronic pain return patients to work? A review and meta-analysis of the literature. *Spine*;**19**:643–52.

Dahl JB, Moiniche S, Kehlet H, 1994. Wound infiltration with local anaesthetics for postoperative pain relief. *Acta Anaesthesiol Scand*;**38**:7–14.

Dahl SL, 1993. Nabumetone: a "nonacidic" nonsteroidal antiinflammatory drug. *Ann Pharmacother*;**27**:456–63.

de Craen AJM, di Giulio G, LampeSchoenmaeckers AJE, Kessels AGH, Kleijnen J, 1996. Analgesic efficacy and safety of paracetamol-codeine combinations versus paracetamol alone: a systematic review. *BMJ*;**313**:321–5.

Devine EC, 1992. Effects of psychoeducational care for adult surgical patients: a meta-analysis of 191 studies. *Patient Educ Couns*;**19**:129–42.

Devine EC, Cook TD, 1986. Clinical and cost saving effects of psychoeducational interventions with surgical patients: a meta-analysis. *Res Nurs Health*;**9**:89–105.

Devine EC, Westlake SK, 1995. The effects of psycho-educational care provided to adults with cancer: meta-analysis of 116 studies. *Oncol Nurs Forum*;**23**:1369–81.

Dexter F, 1994. Analysis of statistical tests to compare doses of analgesics among groups. *Anesthesiology*; **81**:610–15.

di Fabio RP, 1992. Efficacy of manual therapy. *Phys Ther*;**72**:853–64.

di Fabio RP, 1995. Efficacy of comprehensive rehabilitation programs and back school for patients with low back pain: a meta-analysis. *Phys Ther*;**75**:865–78.

Dickersin K, 1989. Pharmacological control of pain during labour. In: Chalmers I, Enkin M, Keirse MJNC, editors. Effective care in pregnancy and childbirth. Oxford: Oxford University Press:913–50.

Dordain G, Aumaitre O, Eschalier A, Decamps A, 1984. La vitamine B12, une vitamine antalgique? Etude critique de la litterature [Vitamin B12, an analgesic vitamin? Critical examination of the literature]. *Acta Neurol Belg*;**84**:5–11.

Downie WW, 1993. Diclofenac/misoprostol. A review of the major clinical trials evaluating its clinical efficacy and upper gastrointestinal tolerability in rheumatoid arthritis and osteoarthritis. *Drugs*;**45**:1–6; discussion 36–7.

Duckro PN, Cantwell-Simmonds E, 1989. A review of studies evaluating biofeedback and relaxation training in the management of pediatric headache. *Headache*;**29**:428–33.

Eandi M, della Pepa C, Rubinetto MP, 1991. Il piroxicam in analgesia [Piroxicam in analgesia]. *Clin Ter*;**136**:107–35.

Edwards JE, McQuay HJ, Moore RA, 1999. Systematic review: dihydrocodeine in postoperative pain. Cochrane Library. In press.

Einarson TR, McGhan WF, Bootman JL, Sabers DL, 1985. Meta-analysis: quantitative integration of independent research results. *Am J Hosp Pharmacy*;**42**:1957–64.

Eisenberg E, Berkey CS, Carr DB, Mosteller F, Chalmers TC, 1994.Efficacy and safety of nonsteroidal anti-inflammatory drugs for cancer pain: a meta-analysis. *J Clin Oncol*;**12**:2756–65.

Eisenberg E, Carr DB, Chalmers TC, 1995. Neurolytic celiac plexus block for treatment of cancer pain: a meta-analysis. *Anesth Analg*;**80**:290–5.

Ernst E, Fialka V, 1993. A review of the clinical effectiveness of exercise therapy for intermittent claudication. *Arch Intern Med*;**153**:2357–60.

Ernst E, Fialka V, 1993. Low-Dose-Lasertherapie: eine kritische Prufung der klinischen Wirksamkeit [Low-dose laser therapy: critical analysis of clinical effect]. *Schweiz Med Wochenschr*;**123**:949–54.

Ernst E, Nielsen GL, Schmidt EB, 1994. Fiskeolie og reumatoid artrit [Fish oils and rheumatoid arthritis]. *Ugeskr Laeger*;**156**:3490–5.

Felson DT, Anderson JJ, Meenan RF, 1990. The comparative efficacy and toxicity of second-line drugs in rheumatoid arthritis. *Arthritis Rheum*;**33**:1449–61.

Felson DT, Anderson JJ, Meenan RF, 1992. Use of short-term efficacy/toxicity tradeoffs to select second-line drugs in rheumatoid arthritis. *Arthritis Rheum*;**35**:1117–25.

Felson DT, Anderson JJ, Boers M, Bombardier C, Chernoff M, Fried B, *et al.*, 1993. The American College of Rheumatology preliminary core set of disease activity measures for rheumatoid arthritis clinical trials. *Arthritis Rheum*;**36**:729–40.

Felson DT, Anderson JJ, Meenan RF, 1994. The efficacy and toxicity of combination therapy in rheumatoid arthritis. A meta-analysis. *Arthritis Rheum*;**37**:1487–91.

Fernandez E, Turk DC, 1989. The utility of cognitive coping strategies for altering pain perception: a meta analysis. *Pain*;**38**:123–35.

Ferraz MB, Atra, 1989. Metaanálise: nova opcao em pesquisa clínica. Principios básicos e um exemplo clínico [Meta-analysis: a new option in clinical research. Basic principles and a clinical example]. *Rev Paul Med*;**107**:5–9.

Ferraz MB, Tugwell P, Goldsmith CH, Atra E, 1990. Meta-analysis of sulfasalazine in ankylosing spondylitis. *J Rheumatol*;**17**:1482–6.

Finley GA, McGrath PJ, Forward SP, McNeill G, Fitzgerald P, 1996. Parents'management of children's pain following 'minor' surgery. *Pain*;**64**:83–7.

Fleet RP, Dupuis G, Marchand A, Burelle D, Beitman BD, 1994. Panic disorder, chest pain and coronary artery disease: literature review. *Can J Cardiol*;**10**:827–34.

Flor H, Fydrich T, Turk DC, 1992. Efficacy of multi-disciplinary pain treatment centers: a meta-analytic review. *Pain*;**49**:221–30.

France RD, Houpt JL, Ellinwood EH, 1984. Therapeutic effects of antidepressants in chronic pain. *Gen Hosp Psychiatry*;**6**:55–63.

Fullerton T, Gengo FM, 1992. Sumatriptan: a selective 5-hydroxytryptamine receptor agonist for the acute treatment of migraine. *Ann Pharmacother*;**26**:800–8.

Furey SA, Wksman JA, Dash BH, 1992. Nonprescription ibuprofen: side effect profile. *Pharmacotherapy*;**12**:403–7.

Gabriel SE, Jaakkimainen L, Bombadier C, 1991. Risk for serious gastrointestinal complications related to use of nonsteroidal anti-inflammatory drugs. A meta-analysis. *Ann Intern Med*;**115**:787–96.

Gabriel SE, Jaakkimainen L, Bombadier C, 1993. The cost-effectiveness of misoprostol for nonsteroidal anti-inflammatory drug-associated adverse gastrointestinal events. *Arthritis Rheum*;**36**:447–59.

Gadsby JG, Flowerdew MW, 1997. The effectiveness of transcutaneous electrical nerve stimulation (TENS) and acupuncture-like transcutaneous electrical nerve stimulation (ALTENS) in the treatment of patients with chronic low back pain. In: Bombardier C, Nachemson A, Deyo R, de Bie R, Bouter L, Shekelle P, *et al.*, editors. CMSG Back Module of the Cochrane Database of Systematic Reviews [updated 5 December 1996]. Oxford: Cochrane Collaboration.

Gam AN, Thorsen H, Lonnberg F, 1993. The effect of low-level laser therapy on musculoskeletal pain: a meta-analysis. *Pain*;**52**:63–6.

Garzillo MJ, Garzillo TA, 1994. Does obesity cause low back pain? *J Manipulative Physiol Ther*;**17**:601–4.

Gebhardt WA, 1994. Effectiveness of training to prevent job-related back pain: a meta-analysis. *Br J Clin Psychol*;**33**:571–4.

Geis GS, 1992. Overall safety of Arthrotec. *Scand J Rheumatol*;suppl 96:33–6.

Genuis ML, 1995. The use of hypnosis in helping cancer patients control anxiety, pain and emesis: a review of recent empirical studies. *Am J Clin Hypn*;**37**:316–25.

Ghent WR, Eskin BA, Low DA, Hill LP, 1993. Iodine replacement in fibrocystic disease of the breast. *Can J Surg*;**36**:453–60.

Glazer S, Portenoy RK, 1991. Systemic local anesthetics in pain control. *J Pain Symptom Manag*;**6**:30–9.

Good M, 1996. Effect of relaxation and music on postoperative pain: a review. *J Adv Nursing*;**24**:905–1014.

Goodkin K, Vrancken MAE, Feaster D, 1995. On the putative efficacy of the antidepressants in chronic, benign pain syndromes. An update. *Pain Forum*;**4**:237–47.

Goodnick PJ, Sandoval R, 1993. Psychotropic treatment of chronic fatigue syndrome and related disorders. *J Clin Psychiatry*;**54**:13–20.

Gøtzsche PC, 1989. Meta analysis of grip strength: most common, but superfluous variable in comparative NSAID trials. *Dan Med Bull*;**36**:493–5.

Gøtzsche PC, 1989. Methodology and overt and hidden bias in reports of 196 double blind trials of nonsteroidal antiinflammatory drugs in rheumatoid arthritis. *Control Clin Trials*;**10**:31–56. [erratum: *Control Clin Trials* 1989;**50**:356].

Gøtzsche PC, 1989. Patients' preference in indomethacin trials: an overview. *Lancet*;**i**:88–91.

Gøtzsche PC, 1989. Review of dose response studies of NSAIDs in rheumatoid arthritis. *Dan Med Bull*;**36**:395–9.

Gøtzsche PC, 1990. Sensitivity of effect variables in rheumatoid arthritis: a meta analysis of 130 placebo controlled NSAID trials. *J Clin Epidemiol*;**43**:1313–18.

Gøtzsche PC, 1993. Meta-analysis of NSAIDs: contribution of drugs, doses, trial designs, and meta-analytic techniques. *Scand J Rheumatol*;**22**:255–60.

Gøtzsche PC, Pødenphant J, Olesen M, Halberg P, 1992. Meta-analysis of second-line antirheumatic drugs: sample bias and uncertain benefit. *J Clin Epidemiol*;**45**:587–94.

Grant A, 1989. The choice of suture materials and techniques for repair of perineal trauma: an overview of the evidence from controlled trials. *Br J Obstet Gynaecol*;**96**:1281–9.

Grant A, Sleep J, 1989. Relief of peritoneal pain and discomfort after childbirth. In Chalmers I, Enkin M, Keirse MJNC, editors. Oxford: Oxford University Press;1347–58.

Haase W, Fischer M, 1991. Statistische metaanalyse von multizentrischen klinischen studien mit ibuprofen im hinblick auf die kohortengrosse [Statistical meta-analysis of multicenter clinical studies of ibuprofen with regard to cohort size]. *Z Rheumatol*;**50**:77–83.

Halpern S, Preston R, 1994. Postdural puncture headache and spinal needle design. Meta-analyses. *Anesthesiology*;**81**:1376–83.

Hanks GW, 1989. Controlled release morphine (MST Contin) in advanced cancer. The European experience. *Cancer*;**63**:2378–82.

Harrison DL, Slack MK, 1996. Meta-analytic review of the effect of subcutaneous sumatriptan in migraine headache. *J Pharm Technol*;**12**:109–14.

Haselkorn JK, Turner JA, Diehr PK, Ciol MA, Deyo RA, 1994. Meta-analysis. A useful tool for the spine researcher. *Spine*;**19**:2076–82S.

Hathaway D, 1986. Effect of preoperative instruction on postoperative outcomes: a meta-analysis. *Nurs Res*;**35**:269–75.

Heller CA, Ingelfinger JA, Goldman P, 1985. Nonsteroidal antiinflammatory drugs and aspirin – analyzing the scores. *Pharmacotherapy*;**5**:30–8.

Hersh EV, Gooper SA, Betts N, Wedell D, MacAfee K, Quinn P, *et al.*, 1993. Single dose and multidose analgesic study of ibuprofen and neclofenamate sodium after third molar surgery. *Oral Surg Oral Med Oral Pathol*;**76**:680–7.

Hill AG, 1967. Review of flufenamic acid in rheumatoid arthritis. *Ann Phys Med*;suppl:87–92.

Hoffman RM, Kent DL, Deyo RA, 1991. Diagnostic accuracy and clinical utility of thermography for lumbar radiculopathy. A meta-analysis. *Spine*;**16**:623–8.

Holmes B, Brogden RN, Heel RC, Speight TM, Avery GS, 1984. Flunarizine. A review of its pharmacodynamic and pharmacokinetic properties and therapeutic use. *Drugs*;**27**:6–44.

Holroyd KA, Penzien DB, 1986. Client variables and the behavioral treatment of recurrent tension headache: a meta analytic review. *J Behav Med*;**9**:515–36.

Holroyd KA, Penzien DB, 1989. Meta-analysis minus the analysis: a prescription for confusion. *Pain*;**39**:359–61.

Holroyd KA, Penzien DB, 1990. Pharmacological versus non pharmacological prophylaxis of recurrent migraine headache: a meta analytic review of clinical trials. *Pain*;**42**:1–13.

Holroyd KA, Penzien DB, Cordingley GE, 1991. Propranolol in the management of recurrent migraine: a meta-analytic review. *Headache*;**31**:333–40.

Holroyd KA, Athens OH, Penzien DB, Jackson MS, Rokicki La, Cordingley GE, 1992. Flunarizine vs propranolol: a meta-analysis of clinical trials. *Headache*;**32**:256.

Howell C, Chalmers I, 1991. A review of prospectively controlled comparisons of epidural with non-epidural forms of pain relief during labour. *Int J Obstet Anesth*;**2**:1–17.

Hurwitz EL, Aker PD, Adams AH, Meeker WC, Shekelle PG, Barr JSJ, 1996. Manipulation and mobilization of the cervical spine: a systematic review of the literature. *Spine*;**21**:1746–60.

Hyman RB, Feldman HR, Harris RB, Levin RF, Malloy GB, 1989. The effects of relaxation training on clinical symptoms: a meta analysis. *Nurs Res*;**38**:216–20.

Jadad AR, 1996. Meta-analysis in pain relief: a valuable but easily misused tool. *Curr Opin Anaesth*;**9**:426–9.

Jadad AR, McQuay HJ, 1996. Meta-analyses to evaluate analgesic interventions: a systematic qualitative review of their methodology. *J Clin Epidemiol*;**49**:235–43.

Jadad AR, Carroll D, Glynn CJ, McQuay HJ, 1995. Intravenous regional sympathetic blockade for pain relief in reflex sympathetic dystrophy: a systematic review and a randomized, double-blind crossover study. *J Pain Symptom Manag*;**10**:13–20.

Jadad AR, Moore RA, Carroll D, Jenkinson C, Reynolds DJM, Gavaghan DJ, *et al.,* 1996. Assessing the quality of reports of randomized clinical trials: is blinding necessary? *Control Clin Trials*;**17**:1–12.

Johanson JF, Rimm A, 1992. Optimal nonsurgical treatment of hemorrhoids: a comparative analysis of infrared coagulation, rubber band ligation and injection sclero-therapy. *Am J Gastroenterol*;**87**:1601–6.

Kaiko RF, Grandy RP, Oshlack B, Pav J, Horodniak J, Thomas G, *et al.,* 1989. The United States experience with oral controlled release morphine (MS Contin tablets). Parts I and II. Review of nine dose titration studies and clinical pharmacology of 15 mg, 30 mg, 60 mg, and 100 mg tablet strengths in normal subjects. *Cancer*;**63**:2348–54.

Kalso E, Tramèr M, Carroll D, McQuay H, Moore RA, 1996. Pain relief from intra-articular morphine after knee surgery: a qualitative systematic review. *Pain*;**71**:642–51.

Kellihan MJ, Mangino PD, 1992. Pamidronate. *Ann Pharmacother*;**26**:1262–9.

Kinnison M, Powe N, Seinberg E, 1989. Results of randomized controlled trials of low versus high osmolality contrast media. *Radiology*;**170**:381–9.

Koes BW, Assendelft WJ, van der Heijden GJ, Bouter LM, Knipschild PG, 1991. Spinal manipulation and mobilisation for back and neck pain: a blinded review. *BMJ*;**303**:1298–303.

Koes BW, Bouter LM, Beckerman H, van der Heijden GJ, Knipschild PG, 1991. Physiotherapy exercises and back pain: a blinded review. *BMJ*;**302**:1572–6.

Koes BW, van Tulder MW, van der Windt WM, Bouter LM, 1994. The efficacy of back schools: a review of randomized clinical trials. *J Clin Epidemiol*;**47**:851–62.

Koes BW, Bouter LM, van der Heijden GJ, 1995. Methodological quality of randomized clinical trials on treatment efficacy in low back pain. *Spine*;**20**:228–35.

Koes BW, Scholten RPM, Mens JMA, Bouter LM, 1995. Efficacy of epidural steroid injections for low-back pain and sciatica: a systematic review of randomized clinical trials. *Pain*;**63**:279–88.

Kubiena G, 1989. [Considerations of the placebo problem in acupuncture. Reflections on usefulness, ethical justification, standardization and differentiated use of placebos in acupuncture]. *Wein Klin Wochenschr*;**101**:362–7.

Labrecque M, Dostaler LP, Rousselle R, Nguyen T, Poirier S, 1994. Efficacy of nonsteroidal anti-inflammatory drugs in the treatment of acute renal colic. A meta-analysis. *Arch Intern Med*;**154**:1381–7.

Ladas SD, Raptis SA, 1992. Conservative treatment of acute pancreatitis: the use of somatostatin. *Hepatogastoenterology*;**39**:466–9.

Lahad A, Malter AD, Berg AO, Deyo RA, 1994. The effectiveness of four interventions for the prevention of low back pain. *JAMA*;**272**:1286–91.

Lancaster T, Silagy C, Gray S, 1995. Primary care management of acute herpes zoster: systematic review of evidence from randomized controlled trials. *Br J Gen Pract*;**45**:39–45.

Laska EM, Sunshine A, Zighelbolm I, Roure C, Marrero I, Wanderling J, *et al.,* 1983. Effect of caffeine on acetaminophen analgesia. *Clin Pharmacol Ther*;**33**:498–509.

Laska EM, Sunshine A, Mueller F, Elvers WB, Siegel C, Rubin A, 1984. Caffeine as an analgesic adjuvant. *JAMA*;**251**:1711–18.

Lataste X, Taylor P, Notter M, 1989. DHE nasal spray in the acute management of migraine attacks. *Cephalalgia*;**9**:342–3.

Laurent MR, Buchanan WW, Bellamy N, 1991. Methods of assessment used in ankylosing spondylitis clinical trials: a review. *Br J Rheumatol*;**30**:326–9.

Letzel H, Schoop W, 1992. Gingko-biloba-Extrakt EGb 761 und Pentoxifyllin bei Claudicatio intermittens. Sekundaranalyse zur klinischen Wirksamkeit [Gingko biloba extract EGb 761 and pentoxifylline in intermittent claudication. Secondary analysis of the clinical effectiveness]. *Vasa*;**21**:403–10.

Lipton RB, Stewart WF, 1994. The epidemiology of migraine. *Eur Neurol*;**34** suppl 2:6–11.

Loosemore TM, Chalmers TC, Dormandy JA, 1994. A meta-analysis of randomized placebo control trials in Fontaine stages III and IV peripheral occlusive arterial disease. *Int Angiol*;**13**:133–42.

Lycka BA, 1990. Postherpetic neuralgia and systemic corticosteroid therapy. Efficacy and safety. *Int J Dermatol*;**29**:523–7.

Macarthur C, Saunders N, Feldman W, 1995. *Helicobacter pylori,* gastroduodenal disease, and recurrent abdominal pain in children. *JAMA*;**273**:729–34.

Magni G, 1991. The use of antidepressants in the treatment of chronic pain. A review of the current evidence. *Drugs*;**42**:730–48.

Mahon WA, de Gregorio M, 1985. Benzydamine: a critical review of clinical data. *Int J Tissue React*;**7**:229–35.

Major P, 1992. Bor ikke-steroide antiinflammatoriske midler brukes ved akutte blotdelsskader? [Would non-steroidal anti-inflammatory agents be used in acute soft tissue injuries?] *Tidsskr Nor Laegeforen*;**112**:1614–15.

Malone MD, Strube MJ, 1988. Meta-analysis of non-medical treatments for chronic pain. *Pain*;**34**:231–44. [erratum: *Pain* 1989;**37**:128].

Malone MD, Strube MJ, Scogin FR, 1989. Reply to Holroyd and Penzien. *Pain*;**39**:362–3.

Max MB, 1995. Thirteen consecutive well-designed randomized trials show that antidepressants reduce pain in diabetic neuropathy and postherpetic neuralgia. *Pain Forum*;**4**:248–53.

McGee J, Alexander M, 1979. Phenothiazine analgesia: fact or fantasy? *Am J Hosp Pharm*;**36**:633–40.

McQuay HJ, 1995. Pre-emptive analgesia: a systematic review of clinical studies. *Ann Med*;**27**:249–56.

McQuay H, Moore RA, 1996. Epidural steroids [letter]. *Anaesth Intensive Care*;**24**:284–6.

McQuay H, Carroll D, Jadad AR, Wiffen P, Moore A, 1995. Anticonvulsant drugs for management of pain: a systematic review. *BMJ*;**311**:1047–52.

McQuay H, Carroll D, Moore A, 1996. Variation in the placebo effect in randomised controlled trials of analgesics: all is as blind as it seems. *Pain*;**64**:331–5.

McQuay H, Tramèr M, Nye BA, Carroll D, Wiffen PJ, Moore RA, 1996. A systematic review of antidepressants in neuropathic pain. *Pain*;**68**:217–27.

McQuay HJ, Carroll D, Moore RA, 1999. Injected morphine in postoperative pain: a quantitative systematic review. *J Pain Symptom Manage*. In press.

McQuay HJ, Carroll D, Moore RA, 1997. Radiotherapy for painful bony metastases: a systematic review. *Clin Oncol*;**42**:71–9.

Mendelson G, 1977. Acupuncture analgesia. I. Review of clinical studies. *Aust N Z J Med*;**7**:642–8.

Mizraji M, 1990. Clinical response to etodolac in the management of pain. *Eur J Rheumatol Inflamm*;**10**:35–43.

Montastruc JL, Senard JM, 1992. Medicaments anticalciques et prophylaxie de la migraine. *Pathol Biol Paris*;**40**:381–8.

Moore RA, McQuay HJ, 1997. Single-patient data meta-analysis of 3453 postoperative patients: oral tramadol versus placebo, codeine and combination analgesics. *Pain*;**69**:287–94.

Moore A, McQuay H, Gavaghan D, 1996. Deriving dichotomous outcome measures from continuous data in randomised controlled trials of analgesics. *Pain*;**66**:229–37.

Moore A, Collins S, Carroll D, McQuay H, 1997. Paracetamol with and without codeine in acute pain: a quantitative systematic review. *Pain*;**70**:193–201.

Moore RA, Collins S, McQuay HJ, 1997. Variation in the placebo effect; the impact on individual trials and consequences for meta-analysis. In: Jensen TS, Hammond DL, Jensen TS, editors. Proceedings of the 8th World Congress on Pain. Progress in Pain Research and Management, vol 2. Seattle: IASP Press:469–76.

Moore RA, Nye BA, Carroll D, Wiffen PJ, Tramèr M, McQuay HJ, 1998. A systematic review of topically-applied non-steroidal anti-inflammatory drugs. *BMJ*;**316**:333–8.

Moote CA, 1996. Ibuprofen arginine in the management of pain. A review. *Clin Drug Invest*;**11**:1–7.

Morton SC, Williams MS, Keeler EB, Gambone JC, Kahn KL, 1994. Effect of epidural analgesia for labor on the Cesarean delivery rate. *Obstet Gynecol*;**83**:1045–52.

Mullen PD, Laville EA, Biddle AK, Lorig K, 1987. Efficacy of psychoeducational interventions on pain, depression, and disability in people with arthritis: a meta analysis. *J Rheumatol*;**14** supp:33–9.

Murphy NG, Zurier RB, 1991. Treatment of rheumatoid arthritis. *Curr Opin Rheumatol*;**3**:441–8.

Naldi L, Zucchi A, Brevi A, Cavalieri d'Oro L, Cainelli T, 1990. Corticosteroids and post-herpetic neuralgia. *Lancet*;**336**:947.

Nicholas JJ,1994. Physical modalities in rheumatological rehabilitation. *Arch Phys Med Rehabil*;**75**:994–1001.

NIH Technology Assessment Panel, 1996. Integration of behavioural and relaxation approaches into the treatment of chronic pain and insomnia. *JAMA*;**176**:313–18.

Olesen J, 1986. Role of calcium entry blockers in the prophylaxis of migraine. *Eur Neurol*;**25** suppl:72–9.

Onghena P, van Houdenhove B, 1992. Antidepressant-induced analgesia in chronic non-malignant pain: a meta-analysis of 39 placebo-controlled studies. *Pain*;**49**:205–19.

Ottenbacher K, di Fabio RP, 1985. Efficacy of spinal manipulation/mobilization therapy. A meta analysis. *Spine*;**10**:833–7.

Pace NL, 1990. Prevention of succinylcholine myalgias: a meta-analysis. *Anesth Analg*;**70**:477–83.

Pace NL, 1993. The best prophylaxis for succinylcholine myalgias: extension of a previous meta-analysis [letter]. *Anesth Analg*;**77**:1080–1.

Patel M, Gutzwiller F, Paccaud F, Marazzi A, 1989. A meta-analysis of acupuncture for chronic pain. *Int J Epidemiol*;**18**:900–6.

Penzien DB, Holroyd KA, 1987. The behavioral treatment of migraine headache: a meta-analytic review of the literature. *Dissertation Abstracts Int*;**47**: No DA8629941.

Penzien DB, Holroyd KA, Hursey KG, Holm JE, Wittchen H, 1987. The behavioral treatment of migraine headache: a meta-analytic review of the literature. [Based on doctoral dissertation by Penzien DB].

Penzien DB, Rains JC, Holroyd KA, 1992. A review of alternative behavioral treatments for headache. *Mis Psychol*;**17**:8–9.

Philipp M, Fickinger M, 1993. Psychotropic drugs in the management of chronic pain syndromes. *Pharmacopsychiatry*;**26**:221–34.

Plantema F, 1986. Worldwide studies comparing piroxicam and naproxen. *Acta Obstet Gynecol Scand Suppl*;138.

Pope JE, Anderson JJ, Felson DT, 1993. A meta-analysis of the effects of nonsteroidal inti-inflammatofy on blood pressure. *Arch Intern Med*;**153**:477–84.

Porzio F, 1993. Meta-analysis of three double-blind comparative trials with sustained-release etodolac in the treatment of osteoarthritis of the knee. *Rheumatol Int*;**13**:S19–24.

Porzio F,1995. Sublingual piroxicam FDDF in the management of pain associated with rheumatic and nonrheumatic conditions; the Italian experience. *Eur J Rheumatol Inflamm*;**15**:11–17.

Post B, Philbrick J, 1988. Do corticosteroids prevent postherpetic neuralgia? A review of the evidence. *J Am Acad Dermatol*;**18**:605–10.

Powell FC, Hanigan WC, Olivero WC, 1993. A risk/benefit analysis of spinal manipulation therapy for relief of lumbar or cervical pain [see comments]. *Neurosurgery*;**33**:73–8; discussion 78–9.

Poynard T, Valterio C, 1994. Meta-analysis of hydroxy-ethylrutosides in the treatment of chronic venous insufficiency. *Vasa*;**23**:244–50.

Pradalier A, Vincent D, 1992. Migraine et anti-inflammatoires non-steroidiens [Migraine and non-steroidal anti-inflammatory agents]. *Pathol Biol Paris*;**40**:397–405.

Puett DW, Griffin MR, 1994. Published trials of nonmedicinal and noninvasive therapies for hip and knee osteoarthritis. *Ann Intern Med*;**121**:133–40.

Pustaver MR, 1994. Mechanical low back pain: etiology and conservative management. *J Manipulative Physiol Ther*;**17**:376–84.

Quan M, 1994. Pelvic inflammatory disease: diagnosis and management. *J Am Board Fam Pract*;**7**:110–23.

Radack K, Deck C, 1987. Do nonsteroidal anti inflammatory drugs interfere with blood pressure control in hypertensive patients? *J Gen Intern Med*;**2**:108–12.

Radack K, Deck C, 1991. Beta-adrenergic blocker therapy does not worsen intermittent claudication in subjects with peripheral arterial disease. A meta-analysis of randomized controlled trials. *Arch Intern Med*;**151**:1769–76.

Radack K, Wyderski RJ, 1990. Conservative management of intermittent claudication. *Ann Intern Med*;**113**:135–46.

Rahimtoola S, 1985. A perspective on the three large multicenter randomized clinical trials of coronary bypass surgery for chronic stable angina. *Circulation*;**72**:V123–35.

Reeve J, Menon D, Corabian P, 1996. Transcutaneous electrical nerve stimulation (TENS): a technology assessment. *Int J Technol Assess Health Care*;**12**:299–324.

Richardson PH, Williams AC, 1993. Meta-analysis of antidepressant-induced analgesia in chronic pain: comment [letter]. *Pain*;**52**:247–9.

Riedemann PJ, Bersinic S, Cuddy LJ, Torrance GW, Tugwell PX, 1993. A study to determine the efficacy of safety and tenoxicam versus piroxicam, diclofenac and indomethacin in patients with osteoarthritis: a meta-analysis. *J Rheumatol*;**20**:2095–103.

Riethmüller-Winzen H, 1987. Flupirtine in the treatment of post-operative pain. *Postgrad Med J*;**63**:61–5.

Robinson M, Mills RJ, Euler AR, 1991. Ranitidine prevents duodenal ulcers associated with non-steroidal anti-inflammatory drug therapy. *Aliment Pharmacol Ther*;**5**:143–50.

Robinson RG, Preston DF, Schiefelbein M, Baxter KG, 1995. Strontium 89 therapy for the palliation of pain due to osseous metastases. *JAMA*;**274**:420–4.

Roderick PJ, Wilkes HC, Meade TW. The gastrointestinal toxicity of aspirin: an overview of randomised controlled trials. *Br J Clin Pharmacol*;**35**:219–26.

Rosch W, 1989. Efficacy of cisapride in the treatment of epigastric pain and concomitant symptoms in non ulcer dyspepsia. *Scand J Gastroenterol*;**24** suppl:165.

Saag KG, Criswell LA, Sems DM, Nettleman MD, Kolluri S, 1996. Low dose corticosteroids in rheumatoid arthritis: a meta-analysis of their moderate-term effectiveness. *Arthritis Rheum*;**39**:1818–25.

Sacks HS, Ancona-Berk VA, Berrier J, Nagalingam R, Chalmers TC, 1988. Dipyridamole in the treatment of angina pectoris: a meta-analysis. *Clin Pharmacol Ther*;**43**:610–15.

Sand T, 1989. Which factors affected reported headache incidences after lumbar myelography? A statistical analysis of publications in the literature. *Neuroradiology*;**31**:55–9.

Schiassi M, Bianchini C, Restelli A, Zoni G, 1989. Tolerability profile of the antiinflammatory compound imidazole salicylate: a metanalysis of safety data in 1408 patients. *Int J Tiss Reac*;**11**:321–6.

Schmader K, Studenski S, 1989. Are current therapies useful for the prevention of postherpetic neuralgia? A critical analysis of the literature. *J Gen Intern Med*;**4**:83–9.

Schuhfried O, Fialka-Moser V, 1995. Iontophoresis in the treatment of pain [in German]. *Wien Med Wochenschr*;**145**:4–8.

Schulz KF, Grimes DA, Altman DG, Hayes RJ, 1996. Blinding and exclusions after allocation in randomised controlled trials: survey of published parallel group trials in obstetrics and gynaecology. *BMJ*;**312**:742–4.

Shekelle PG, Adams AH, Chassin MR, Hurwitz EL, Brook RH, 1992. Spinal manipulation for low-back pain. *Ann Intern Med*;**117**:590–8.

Siminoski K, Josse RG, Bowyer M, 1996. Calcitonin in the treatment of osteoporosis. *Can Med Assoc J*;**155**:962–5.

Simkin P, 1989.Non-pharmacological methods of pain relief during labour. In: Chalmers I, Enkin M, Keirse MJNC, editors. Effective care in pregnancy and childbirth. Oxford: Oxford University Press:899–911.

Simmons J, Stavinoha W, Knodel L, 1984. Update and review of chemonucleolysis. *Clin Orthop*;Mar:51–60.

Sindhu F, 1996. Are non-pharmacological nursing interventions for the management of pain effective? A meta-analysis. *J Adv Nurs*;**24**:1152–9.

Smith MC, Holcombe JK, Stullenbarger E, 1994. A meta-analysis of intervention effectiveness for symptom management in oncology nursing research. *Oncol Nurs Forum*;**21**:1201–9; discussion 1209–10.

Solomon GD, 1989. Verapamil in migraine prophylaxis – a five year review. *Headache*;**29**:425–7.

Spierings EL, Messinger HB, 1988. Flunarizine vs. pizotifen in migraine prophylaxis: a review of comparative studies. *Cephalalgia*;**8** suppl:27–30.

Stalnikowicz R, Rachmilewitz D, 1993. NSAID-induced gastroduodenal damage: is prevention needed? a review and meta-analysis. *J Clin Gastroenterol*;**17**:238–43.

Stewart WF, Simon D, Shechter A, Lipton RB, 1995. Population variation in migraine prevalence: a meta-analysis. *J Clin Epidemiol*;**48**:269–80.

Stulberg B, Bauer T, Belhobek G, 1990. Making core decompression work. *Clin Orthop*;Dec:186–95.

Suls J, Fletcher B, 1985. The relative efficacy of avoidant and nonavoidant coping strategies: a meta analysis. *Health Psychol*;**4**:249–88.

Suls J, Wan CK, 1989. Effects of sensory and procedural information on coping with stressful medical procedures and pain: a meta-analysis. *J Consult Clin Psychol*;**57**:372–9.

Sutters KA, Miaskowski C, 1992. The problem of pain in children with cancer: a research review. *Oncol Nurs Forum*;**19**:465–71.

Teasell RW, Harth M, 1996. Functional restoration: returning patients with chronic low back pain. *Spine*;**21**:844–7.

ter Riet G, Kleijnen J, Knipschild P, 1990. Acupuncture and chronic pain: a criteria-based meta-analysis. *J Clin Epidemiol*;**43**:1191–9.

Tfelt Hansen P, 1986. Efficacy of beta blockers in migraine. A critical review. *Cephalalgia*;**6** suppl:15–24.

Tfelt-Hansen P, 1993. Sumatriptan for the treatment of migraine attacks – a review of controlled clinical studies. *Cephalalgia*;**13**:238–44.

Thomas D, Williams RA, Smith DS, 1980. The frozen shoulder: a review of manipulative treatment. *Rheumatol Rehab*;**19**:173–9.

Thomson AM, Hillier VF, 1994. A re-evaluation of the effect of pethidine on the length of labour. *J Adv Nurs*;**19**:448–56.

Trabant H, Widdra W, De LS, 1990. Efficacy and safety of intranasal buserelin acetate in the treatment of endometriosis: a review of six clinical trials and comparison with danazol. *Prog Clin Biol Res*;**323**:357–82.

Tramèr M, Williams J, Carroll D, Wiffen PJ, McQuay HJ, Moore RA, 1997. Systematic review of direct comparisons of non-steroidal anti-inflammatory drugs given by different routes for acute pain. *Acta Anaesth Scand*;**42**:71–9.

Turner JA, 1995. Efficacy of antidepressant medication for chronic pain problems. *Pain Forum*;**4**:254–5.

Turner JA, Denny MC, 1993. Do antidepressant medications relieve chronic low back pain? *J Fam Pract*;**37**:545–53.

Turner JA, Ersek M, Herron L, Deyo R, 1992. Surgery for lumbar spinal stenosis. Attempted meta-analysis of the literature. *Spine*;**17**:1–8.

Turner JA, Deyo RA, Loeser JD, von Korff M, Fordyce WE, 1994. The importance of placebo effects in pain treatment and research. *JAMA*;**271**:1609–14.

Turner JA, Loeser JD, Bell KG, 1995. Spinal cord stimulation for chronic low back pain; a systematic literature synthesis. *Neurosurgery*;**37**:1088–96.

Vaiani G, Grossi E, 1990. Meta-analysis of italian clinical trials of nabumetone. *Drugs*;**40**:48–9.

van der Heihden CJM, Beurskens AJ, Koes BW, Assendelft WJ, de Vet HC, Bouter LM, 1995. The efficacy of traction for back and neck pain: a systematic, blinded review of randomized clinical trial methods. *Phys Ther*;**75**:93–104.

van der Heijden CJM, van der Windt DAW, Kleijnen J, Koes BW, Bouter LM, 1996. Steroid injections for shoulder disorders: a systematic review of randomized clinical trials. *Br J Gen Pract*;**46**:309–16.

van der Hoogen HM, Koes BW, van Eijk JT, Bouter LM, 1995. On the accuracy of history, physical examination, and erythrocyte sedimentation rate in diagnosing low back pain in general practice. A criteria-based review of the literature. *Spine*;**20**:318–27.

Vischer T, 1987. Efficacy and tolerability of tenoxicam: an overview. *Eur J Rheumatol Inflamm*;**9**:51–7.

Vischer T, 1987. Efficacy and tolerability of tenoxicam: an overview. *Scand J Rheumatol Suppl*;**65**:107–12.

Volmink J, Lancaster T, Gray S, Silagy C, 1996. Treatments for postherpetic neuralgia – a systematic review of randomized controlled trials. *Fam Pract*;**13**:84–91.

Watts RW, Silagy CA, 1995. A meta-analysis on the efficacy of epidural corticosteroids in the treatment of sciatica. *Anaesth Intensive Care*;**23**:564–9.

White KP, Harth M, 1996. An analytical review of 24 controlled clinical trials for fibromyalgia syndrome (FMS). *Pain*;**64**:211–19.

Wilkie DJ, Savedra MC, Holzemer WL, Tesler MD, Paul SM, 1990. Use of the McGill Pain Questionnaire to measure pain: a meta-analysis. *Nurs Res*;**39**:36–41.

Windle M, Booker L, Rayburn W, 1989. Postpartum pain after vaginal delivery. A review of comparative analgesic trials. *J Reprod Med*;**34**:891–5.

Wood MJ, 1996. How to measure and reduce the burden of zoster-associated pain. *Scand J Infect Dis*;suppl:55–8.

Wood MJ, Kay R, Dworkin RH, Soong SJ, Whitley RJ, 1996. Oral acyclovir therapy accelerates pain resolution in patients with herpes zoster: a meta-analysis of placebo-controlled trials. *Clin Infect Dis*;**22**:341–7.

Wright JM, 1994. Review of the symptomatic treatment of diabetic neuropathy. *Pharmacotherapy*;**14**:689–97.

Yee LY, Lopez JR, 1992. Transdermal fentanyl. *Ann Pharmacother*;**26**:1393–9.

Zhang WY, Li WPA, 1994. The effectiveness of topically applied capsaicin. A meta-analysis. *Eur J Clin Pharmacol*;**46**:517–22.

Zhang WY, Li Wan Po A, 1996. Analgesic efficacy of paracetamol and its combination with codeine and caffeine in surgical pain: a meta-analysis. *J Clin Pharm Ther*;**21**:261–82.

Chapter 8

Paracetamol with and without codeine in acute pain

Summary

A systematic review of RCTs to assess the analgesia obtained from single oral doses of paracetamol alone and in combination with codeine in postoperative pain found:

- 39 trials of paracetamol against placebo with 4124 patients
- 21 trials of paracetamol plus codeine against placebo with 450 patients
- 12 trials of paracetamol plus codeine against the same dose of paracetamol with 794 patients.

Pain relief information was extracted and converted into dichotomous information, that is, numbers of patients with at least 50% pain relief. Wide variation in responses to placebo (0–72%) and active drugs (5–89%) were observed.

In postoperative pain states paracetamol, 1000 mg, alone against placebo had an NNT of 4.6 (95% CI, 3.9–5.4) and paracetamol, 600/650 mg, alone an NNT of 5.3 (95% CI, 4.1–7.2). Paracetamol, 600/650 mg, plus codeine, 60 mg, against placebo had a better NNT of 3.1 (95% CI, 2.6–3.9), with no overlap of 95% CIs with paracetamol, 600/650 mg, alone. In direct comparisons of paracetamol plus codeine with paracetamol alone, the additional analgesic effect of 60 mg of codeine added to paracetamol was 11 extra patients in every 100 achieving at least 50% pain relief. In indirect comparisons of each with placebo, it was 14 extra patients per 100. This gave an NNT for adding codeine, 60 mg, of 7.7 (95% CI, 5.1–20).

The results confirm that paracetamol is an effective analgesic and that codeine, 60 mg, added to paracetamol produces worthwhile additional pain relief even in single oral doses.

This systematic review has been published in part by Moore and colleagues.[1]

Introduction

Paracetamol is an important non-opiate analgesic, which is commonly prescribed as well as being available without prescription. In England in 1995, paracetamol alone accounted for over 5 million prescriptions for adults (16% of total non-opiate analgesic prescriptions), with 4.5 million prescriptions of paediatric suspensions.[2] In combination with codeine, paracetamol accounted for a further 6.4 million prescriptions (20% of total non-opiate analgesics). Paracetamol alone and in combination with a variety of opioids accounted for 93% of prescriptions in this BNF classification.

Policy decisions and guidelines are increasingly being made on the basis of hard evidence. Trying to judge the relative efficacy of analgesics against one another is not easy because there are few such direct comparisons. Only five direct comparisons were found of paracetamol, 1000 mg, and ibuprofen, 400 mg, in acute pain.[3]

Relative efficacy can also be determined indirectly, from comparisons of each analgesic with placebo, using a common descriptor of efficacy, and then comparing the results for various analgesic interventions, both pharmacological and non-pharmacological. In this review of paracetamol, its analgesic efficacy is compared with information about other drugs, again determined by similar quantitative systematic reviews.

Methods

RCTs of paracetamol in postoperative pain (post dental extraction, postsurgical or postpartum pain) were sought. A number of different search strategies were used to identify eligible reports in MEDLINE (1966–May 1996), EMBASE (1980– 96), Cochrane Library (March 1996) and the Oxford Pain Relief database (1950–94).[4] The words 'paracetamol', 'acetaminophen', and 'trial' were used in a free text search, both alone and in combination, and without restriction to language. Additional reports were identified from reference lists of retrieved reports, review articles (including a recent systematic review of paracetamol plus codeine)[5] and textbooks.

Inclusion criteria for paracetamol

Neither pharmaceutical companies nor authors of papers were contacted for unpublished reports. Abstracts and review articles were not considered. The inclusion criteria used were:

- randomised allocation to treatment groups which compared either paracetamol or a paracetamol and codeine combination with placebo or a paracetamol and codeine combination with the same dose of paracetamol alone
- full journal publication
- established postoperative pain, with the pain outcome measured using a five-point pain relief scale with standard wording (none, slight, moderate, good, complete) or a four-point pain intensity scale (none, mild, moderate or severe) or a VAS for pain relief or pain intensity, TOTPAR or SPID (at 4, 5 or 6 hours) as a derived pain relief outcome (or sufficient data provided to allow their calculation)
- postoperative oral administration
- adult patients
- baseline pain of moderate to severe intensity (for VAS this equates to > 30 mm)[6]
- double-blind design.

Reports for the relief of other pain conditions were excluded, as were those for paracetamol used in combination with drugs other than codeine and trials where the number of patients per treatment group was less than ten.[7] In postpartum pain, trials were included if the pain investigated resulted from episiotomy or Caesarean section combined with uterine cramps but trials investigating uterine cramps alone were excluded.

Data extraction and analysis

The numbers of patients treated, the mean TOTPAR, SPID, VAS TOTPAR or VAS SPID, study duration and the dose given were taken from each report. Information on adverse events was also extracted. For each report, the mean TOTPAR, SPID, VAS TOTPAR or VAS SPID values for active and placebo were converted to % maxTOTPAR or % maxSPID by division into the calculated maximum value.[8] The proportion of patients in each treatment group who achieved at least 50% maxTOTPAR was calculated using verified equations.[9–11] These proportions were then converted into the number of patients achieving at least 50% maxTOTPAR by multiplying by the total number of patients in the treatment group. Information on the number of patients with at least 50% maxTOTPAR for treatment and placebo was then used to calculate relative benefit and NNT.

Relative benefits estimates were calculated with 95% CIs using a random effects model;[12] the random effects model was chosen because it produces the most conservative estimate (homogeneity was assumed when $p > 0.1$). NNTs and 95% CIs were calculated using the method described by Cook and Sackett.[13] A statistically significant difference from control was assumed when the 95% CI of the relative benefit did not include 1.

Results

Paracetamol versus placebo

The literature searches found 37 reports of 39 trials which fulfilled the inclusion criteria; 2530 patients were given paracetamol and 1594 patients placebo. Details of the trials are presented in *Table 23*.

Of the trials found, 21 investigated oral surgery pain (post-dental pain, predominantly third molar extraction with bone removal), eight postsurgical pain (elective general, gynaecological and orthopaedic surgery) and ten postpartum pain (episiotomy and post-Caesarean section). The doses of paracetamol administered were 500 mg in six trials, 600 mg in six trials, 650 mg in 11 trials and 1000 mg in 20 trials; for analysis purposes, data from paracetamol, 600 mg, and paracetamol, 650 mg, were combined. One report on episiotomy provided dichotomous information on the overall patient global rating of pain relief.[14] This report was included. The proportion of patients with good or excellent pain relief was used.

The variation in placebo response rates (i.e. the proportion of patients with at least 50% pain relief) was from 0% to 72% of patients with at least 50% maxTOTPAR (see *Figure 24*, page 69). The placebo response rate ranged from 0% to 72% in post-dental pain, from 11% to 48% in postsurgical pain, and from 0% to 34% in postpartum pain. The variation in response rates with all doses of paracetamol was 5–83% *(Figure 24)*. The mean response rate for paracetamol, 600/650 mg, was 42% and for placebo 23%.

Combining data across conditions, the pooled relative benefits for all doses of paracetamol versus placebo were significant *(Table 24)*. Paracetamol, 600/650 mg, compared with placebo in single dose administration had an NNT for at least 50% pain relief of 5.3 (95% CI, 4.1–7.2) and at 1000 mg the NNT was 4.6 (95% CI, 3.9–5.4), with overlap between the CIs.

TABLE 23 *Trials of paracetamol versus placebo*

Study	Condition and number of patients	Design, study duration and follow-up	Outcome measures	Dosing regimen	Analgesic outcome results	Remedication	Withdrawals and exclusions	Adverse effects	Quality score
Beaver & McMillan, 1980	Episiotomy and uterine cramp (vaginal delivery < 48 hours). n = 108 Age ?	RCT, double-blind, single oral dose, 5 parallel groups. 3 hour washout prior to start and medication given > 30 minutes before or > 2 hours after patient's meal. Evaluated in hospital by nurse observer at 0 hours then hourly for 6 hours or until pain returned to pre-medication level.	PI (4-point scale) PR (5-point scale) 50% PR (y/n) Global rating (5-point scale) NB: episiotomy pain assessed as 'right now' and uterine cramp pain as 'during the last hour'.	Placebo (n = 22); paracetamol, 1000 mg (n = 22).	Paracetamol significantly superior to placebo ($p < 0.05$–0.01) for all measures except those based on change in PI.	Patients allowed remedication after 2 hours if pain returned to pre-medication levels. After remedication, PR = 0 for all further time points.	108 analysed. 'There were no drop-outs.'	NSD between groups. All mild and subjective. Total numbers reporting adverse effects (number of effects): placebo 7/22 (8); paracetamol 11/22 (13).	4
Bentley & Head, 1987	Third molar bony impacted. n = 128 Mean age, 'mid-20s'	RCT, double-blind, single oral dose, four parallel groups. No information on anaesthesia except 'no sedative or narcotic agents were used before, during or after surgery'. Self-assessed at home at 0, 1, 2, 3, 4, 5 hours then reports posted to investigator.	PI (10-point scale) PR (5-point scale)	Placebo (n = 17); paracetamol, 1000 mg (n = 41).	Paracetamol significantly superior for all measures of efficacy.	Rescue analgesic, Tylenol no 3. For patients with remedication at < 5 hours, last PI and PR scores carried on for all further time points.	120 patients analysed. Exclusions: 3 took no medication; 1 took only part; 1 took none until day after surgery; 1 remedicated after 30 minutes; 1 vomited within 30 minutes of taking medication; 1 did not return forms.	53 patients reported one or more; 86 reported in total, majority being dizziness, drowsiness, nausea and vomiting. NSD between treatment groups. Total numbers reporting adverse effects (number of effects): paracetamol 21/42 (31); placebo 9/19 (16).	3
Berry, *et al.*, 1975	Episiotomy n = 225 Age 15+ years.	RCT, double-blind, single oral dose, 3 parallel groups. 12 hour washout prior to start. Assessed by observer(s) in hospital at 0, 0.5, 1, 2, 3, 4 hours.	PI (5-point scale) PR (5-point scale, but non-standard) Gastric discomfort (4-point scale) Global rating (5-word scale)	Placebo (n = 76); paracetamol, 1000 mg (n = 76).	Total PI and PR scores showed paracetamol significantly superior to placebo. Global rating > good. Placebo 18/76. Paracetamol 43/76.	After reasonable period, rescue analgesia could be prescribed at the investigator's discretion and patient regarded as a treatment failure (no information on how data handled). Placebo 23/76; paracetamol 2/76.	No details.	"None of the patients experienced adverse drug reactions." Number reporting gastric discomfort: paracetamol (slight) 2/76; placebo (moderate) 2/76.	3
Cooper & Beaver, 1976	Impacted third molar n = 216 Age 16+ years	RCT, double-blind, 2 single oral dose studies (paracetamol = study 2), 4 parallel groups. No information on anaesthesia. Self-assessed at home at 0, 1, 2, 3 hours (questionnaire) – posted to investigator.	PI (4-point scale) PR (5-point scale) Pain half gone? (y/n) Had the patient fallen asleep during the hour? (y/n)	Study 2: Placebo (n = 40); paracetamol, 600 mg (n = 40).	Parametric analysis concluded that paracetamol significantly superior to placebo ($p < 0.01$).	Remedication consisted of taking second envelope and evaluating it in same way as the first. Unclear how data was then handled. 112/160 patients required remedication.	160 analysed. Exclusions: 30 patients did not return forms; 12 completed forms improperly; 4 took concomitant medication; 6 required no medication; 4 were randomly deleted to even out numbers on each treatment.	None serious reported. Most common: drowsiness, nausea and headache. Numbers reporting adverse effects (number of effects): placebo 5/40 (5); paracetamol 5/40 (7).	5
Cooper, *et al.*, 1980	Impacted third molar n = 298 Mean age 'early 20s'	RCT, double-blind, single oral dose, 6 parallel groups. No information on anaesthesia. Self-assessed at home at 0, 1, 2, 3, 4 hours (questionnaire) – majority collected 1 week later by observer, few returned by mail (patients telephoned if problems encountered)	PI (4-point scale) PR (5-point scale) Global rating (5-point scale) 50% relief of baseline pain (y/n) Time to remedication	Placebo (n = 38); paracetamol, 500 mg (n = 37).	Paracetamol showed significant analgesic efficacy for all measures.	t > 1 hour before remedication (unclear how data handled). Remedication < 4 hours: placebo 26/38; paracetamol 25/37.	247 analysed. Exclusions: 21 lost to follow-up; 10 dropped out before ingesting medication (no details); 20 ingested medication but excluded for protocol violations (no details).	Numbers reporting adverse events (number of effects): placebo 6/38 (7); paracetamol 3/37 (6).	4

continued

TABLE 23 contd *Trials of paracetamol versus placebo*

Study	Condition and number of patients	Design, study duration and follow-up	Outcome measures	Dosing regimen	Analgesic outcome results	Remedication	Withdrawals and exclusions	Adverse effects	Quality score
Cooper, *et al.*, 1981	Impacted third molar n = 248 Mean age 'early 20s'	RCT, double-blind, single oral dose, 5 parallel groups. Either general or local anaesthetic. Self-assessed at home at 0, 1, 2, 3, 4 hours (questionnaire).	PI (4-point scale) PR (5-point scale) Global rating (5-point scale) 50% relief of baseline pain (y/n) Time to remedication	Placebo (n = 37); paracetamol, 650 mg (n = 37).	All active treatments significantly superior to placebo for all measures.(For mean values of SPID, peak PID, total PR, etc. see Table 2).	Remedication after hour 1 if needed (last score used for all further time points). Remedication < 4 hours: placebo 20/37; paracetamol 2/37.	200 analysed. Exclusions: 17 did not ingest medication; 31 ingested medication but violated protocol (remedicated before 1st hour observation, constant deviation of more than 15 minutes from evaluation times, did not return questionnaire, lost to follow-up)	None serious reported. Numbers reporting adverse effects (number of effects): placebo 4/37 (5); paracetamol 12/37 (15).	4
Cooper, *et al.*, 1986	Oral surgery (involving bone removal) n = 112 Age 16+ years	RCT, double-blind, single oral dose, 3 parallel groups, single centre and 1 surgeon. No information on anaesthesia. Self-assessed at home at 0, 1, 2, 3, 4, 5, 6 hours (diary).	PI (4-point scale) PR (5-point scale) Pain half gone? (y/n) Global rating (5-point scale) Time to remedication	Placebo (n = 22); paracetamol, 1000 mg (n = 38).	For all measures paracetamol significantly superior to placebo ($p < 0.05$). (see Table 2 for mean values of SPID, TOTPAR, 50% reduction, etc.)	Remedication after first hour if needed (last score used for all further time points).	99 analysed. Exclusions: 6 did not require analgesia; 3 fell asleep; other 4 were for 'various protocol violations'.	None serious reported. Over half those reported were drowsiness; there were two reports of nausea. Numbers of events reported: paracetamol 12; placebo 0.	3
Cooper, *et al.*, 1988	Impacted third molar n = 165 Age range 18–57 years	RCT, double-blind, single oral dose, 3 parallel groups, single centre and 1 surgeon. Local anaesthetic with sedative and/or nitrous oxide. 4-hour washout prior to start. Self-assessed at home at 0, 0.5, 1, 2, 3, 4, 5, 6 hours (diary).	PI (4-point scale) PR (5-point scale) Pain half gone? (y/n) Global rating (5-point scale) Time to remedication	Placebo (n = 40); paracetamol, 600 mg (n = 36).	Paracetamol appeared clinically more effective than placebo but was not significantly superior.	Remedication after first hour (last or baseline score used for all further time points).	143 analysed. Exclusions: 11 lost to follow-up; 8 did not require medication; 3 for 'various protocol violations'.	None serious reported. Numbers of events reported: placebo 3; paracetamol 8.	4
Cooper, *et al.*, 1989	Removal of impacted teeth n = 194 Age 16+ years	RCT, double-blind, single oral dose, 3 parallel groups. local anaesthetic with iv sed., atropine and/or nitrous oxide. 4-hour washout prior to start. Self-assessed at home at 0, 0.5, 1, 2, 3, 4, 5, 6 hours (diary).	PI (4-point scale) PR (5-point scale) Pain half gone? (y/n) Global rating (5-point scale) Time to remedication	Placebo (n = 64); paracetamol, 1000 mg (n = 59).	Paracetamol significantly superior to placebo ($p < 0.05$–< 0.001) for all measures. (see Table 2 for mean values, SPID, TOTPAR, etc.)	Remedication after first hour (last or baseline score used for all further time points). Approximately 50% paracetamol and 36% placebo completed least 4 hours before remedication. Significantly longer mean time to remedication for paracetamol ($p < 0.05$) than for placebo.	184 analysed. Exclusions: 4 slept through more than two observations; 2 lost to follow-up; 2 did not need medication; 1 had inadequate baseline PI; 1 failed to complete evaluations at set times.	190 evaluated (all who ingested medication). None serious reported – drowsiness being most common. Numbers reporting adverse effects (number of effects): placebo 7/64 (7); paracetamol 11/63 (13).	5
Cooper & Kupperman, 1991	Removal of one or more impacted teeth n = 247 Age 'young adult'	RCT, double-blind, single oral dose, 6 parallel groups. Local anaesthetic (lidocaine + epinephrine) with i.v. diazepam and methohexital (nitrous oxide also used on occasion). Self-assessed at home at 0, 0.5 hours then hourly for 6 hours (diaries).	PI (4-point scale) PR (5-point scale) Global rating (5-point scale) Time to remedication	Placebo (n = 44); paracetamol, 650 mg (n = 37).	Paracetamol was only active drug not significantly superior to placebo for any measure.	t > 1 hour before remedication, data included and baseline or last score (most severe) used for all further time points.	226 analysed. Exclusions; 13 did not require medication; 3 lost to follow-up, 2 remedicated. with slight pain before second hour observation; 2 remedicated before first hour observation; 1 fell asleep for over 2 hours.	All adverse effects mild. Numbers reporting adverse effects (number of effects): placebo 7/44 (9); paracetamol 6/37 (7).	3

continued

TABLE 23 contd *Trials of paracetamol versus placebo*

Study	Condition and number of patients	Design, study duration and follow-up	Outcome measures	Dosing regimen	Analgesic outcome results	Remedication	Withdrawals and exclusions	Adverse effects	Quality score
Dionne, *et al.*, 1994	Impacted third molar n = 135 Age 16+ years	RCT, double-blind, single oral dose, 5 parallel groups. General and local anaesthetic. 4-hour washout prior to start. Self-assessed at clinic for at least first 2 hours then at home hourly for 6 hours.	PI (4-point scale) PR (5-point scale) Global rating (5-point scale) Time to remedication	Placebo (n = 25); paracetamol, 650 mg (n = 27).	Paracetamol not significantly different to placebo for any measure of analgesia.	t ≥ 2 hours before remedication, data included and baseline used for all further time points.	124 analysed. Exclusions: 4 previously enrolled in study; 3 remedicated before 2 hours; 2 lost to follow-up; 1 ineligible because of codeine sensitivity.	All reported mild. Numbers reporting adverse effects (number of effects): placebo 5/25 (5); paracetamol 7/27 (9).	3
Dolci, *et al.*, 1994	Removal of single impacted third molar n = 336 Age 18+ years	RCT, double-blind, single oral dose, 4 parallel groups, multicentre (11). No information on anaesthesia. 24-hour washout prior to start. Evaluations made at 0, 0.5, 1, 1.5, 2 hours in clinic then at 3 and 4 hours at home (diary).	PI (4-point scale) PR (5-point scale) Global rating (5-point scale)	Placebo (n = 76); paracetamol, 500 mg (n = 72).	PID/PR: paracetamol significantly superior to placebo at *t* = 30 minutes ($p < 0.01$) and at t = 1–4 hours ($p < 0.0001$). SPID/TOTPAR: paracetamol significantly superior at t = 1–4 hours. Dichotomous data available for Global Rating (see page 190, Table V).	Rescue analgesia permitted after 90 minutes, no further evaluation post-remedication. Number remedicated at t > 1.5 hours: placebo 46/76; paracetamol 15/72.	298 analysed. Exclusions: 15 lost to follow-up, 6 remedicated before 1.5 hours; 3 experienced adverse event and did not complete assessment; 14 did not experience > moderate baseline pain.	4 withdrew from study because of adverse effects. Paracetamol: nausea 1; swelling 1. Placebo: fever 1, nausea and diarrhoea 1. Numbers reporting adverse effects (number of effects): paracetamol 7/80 (7); placebo 8/82 (12).	4
Fassolt & Stocker, 1983	Postoperative ('simple surgery' with 15+ surgical techniques named) n = 146 Age 18+ years	RCT, double-blind, single oral dose. 5 parallel groups. General anaesthetic used. 4-hour washout prior to start. Evaluation at 0, 30 minutes then hourly for 6 hours by same trained observer in hospital.	PI (5-point scale and VAS) PR (5-point scale) Global rating (5-point scale) Time to remedication	Placebo (n = 28); paracetamol, 650 mg (n = 29).	All active drugs significantly superior for all measures of efficacy (except Suprofen 200 which showed no significant difference in global rating).	Remedication allowed after 2 hours (last score used for all further time points). Number not remedicated at 6 hours: placebo 9/28; paracetamol 25/29.	No information given.	None serious reported – no further details given.	2
Forbes, *et al.*, 1982	Impacted third molar (1 or more) n = 177 Age 15+ years	Double-blind, single oral dose, 5 parallel groups. No information on anaesthesia. Self-assessed at home at 0, 1 hours, then hourly for 12 hours.	PI (4-point scale) PR (5-point scale) 50% PR? (y/n) Global rating (5-point scale) Time to remedication	Placebo (n = 30); paracetamol, 600 mg (n = 34).	At 4 hours paracetamol significantly superior to placebo for PR, peak PR and 50% PR ($p < 0.01$).	If remedication at < 2 hours -excluded. If > 2 < 12 hours, PR = 0 and PI baseline or last. % patients remedicated at > 2 hours but < 12 hours: placebo 93%; paracetamol 85%.	159 analysed. Exclusions: 4 lost to follow-up; 3 did not take medication; 7 took remedication < 2 hours; 1 medicated with mild pain; 3 did not complete forms properly.	No active treatment produced more than placebo. None were serious.	3
Forbes, *et al.*, 1983	Postoperative (general, gynaecological or orthopaedic surgery) n = 132 Age 18+ years	RCT, double-blind, single oral dose, 5 parallel groups. General anaesthetic. Oral analgesia given on request first day able to take. Self-assessed in hospital at 0, 0.5, 1, 1.5, 2 hours then hourly for 12 hours. Single nurse observer present for first 6 hours, 2nd 6-hour period monitored by ward staff.	PI (4-point scale) PR (5-point scale) Pain half gone? (y/n) Global rating (5-point scale) Time to remedication	Placebo (n = 26); paracetamol, 600 mg (n = 26).	At 6 hours paracetamol significantly superior for all measures; at 12 hours for all except SPID.	Patients remedicated on demand if < 2 hours excluded. If remedicated at < 12 hours, PR = 0, PI = baseline or last score for remaining time points.	132 analysed. There were no exclusions.	None serious reported. Most common were drowsiness, dizziness and dry mouth. Numbers reporting adverse effects (number of effects): placebo 4/26 (5); paracetamol 11/26 (11).	5
Forbes, *et al.*, 1984	Postoperative (general, gynaecological or orthopaedic surgery) n = 132 Age 18+ years	RCT, double-blind, single oral dose, 5 parallel groups. General anaesthetic. Trial drug given on request first day after surgery. Assessed in hospital by single nurse observer at 0, 15 and 30 minutes, then hourly for 6 hours.	PI (4-point scale) PR (5-point scale) Pain half gone? (y/n) Acceptability (5-point scale, each hour) Time to remedication	Placebo (n = 33); paracetamol, 650 mg (n = 31).	Paracetamol significantly superior for all measures of efficacy at t = 1–4 hours. But only marginally significant for peak PI difference and peak PR.	If remedication at < 6 hours, PR = 0, PI = baseline or last for all remaining time points.	129 analysed. Exclusions: 2 remedicated < 2 hours; 1 received interfering medication.	None serious recorded. Sedation accounted for 65–70%. Numbers reporting adverse effects (number of effects): placebo 8/33 (9); paracetamol 9/33 (10).	4

continued

TABLE 23 contd *Trials of paracetamol versus placebo*

Study	Condition and number of patients	Design, study duration and follow-up	Outcome measures	Dosing regimen	Analgesic outcome results	Remedication	Withdrawals and exclusions	Adverse effects	Quality score
Forbes *et al.*, 1984	Impacted third molar (1 or more) n = 191 Age 15+ years	RCT, double-blind, single oral dose, 4 parallel groups. General/local anaesthetic (unclear). Self-assessed at home at 0, 1 hours, then hourly for 6 hours. Returned 5 days later for review and debriefing.	PI (4-point scale) PR (5-point scale) Pain half gone? (y/n) Global rating (5-point scale) Time to remedication	Placebo (n = 36); paracetamol, 650 mg (n = 39).	Paracetamol significantly superior for all measures of total and peak analgesia.	Patients could remedicate after 2 hours but were asked to complete next evaluation before doing so (PI last or baseline score; PR = 0 for all further time points). % remedicating by hour 6: paracetamol 74%; placebo 97%.	148 analysed. Exclusions: 1 did not return results; 1 lost to follow-up; 26 did not require medication, 8 did not follow instructions; 7 remedicated < 2 hours.	NSD between treatments, none serious. Most frequent was drowsiness. Numbers reporting adverse effects (number of effects): placebo 2/40 (2); paracetamol 1/43 (2).	5
Forbes, *et al.*, 1989	Impacted third molar (1 or more) n = 107 Age 15+ years	RCT, double-blind, single oral dose, 4 parallel groups at 2 centres. General and local anaesthetic (unclear). Patients self-assessed at home at 0, 1, 2 hours then hourly for 12 hours or until remedication (diary). Returned 5 days later for review and debriefing.	PI (4-point scale) PR (5-point scale) Global rating (5-point scale) Time to remedication	Placebo (n = 23); paracetamol, 600 mg (n = 22).	For 12 hours: paracetamol not significantly superior for any measure. For 4 hours: paracetamol significantly superior for SPID, TOTPAR and hours of 50% relief.	Patients could remedicate after 2 hours but were asked to complete next evaluation before doing so (PI last or baseline score; PR = 0 for all further time points). % remedicating by hour 12: paracetamol 95%; placebo 91%.	88 analysed. Exclusions: 9 did not take medication; 2 remedicated before 2 hour point; 1 remedicated with slight pain; 4 did not complete evaluation; 1 took only part of the medication; 2 remedicated despite having some relief from the study medication.	None serious reported. Numbers reporting adverse effects (number of effects): placebo 2/26 (2); paracetamol 3/26 (3). NB: includes those reported post-remedication.	
Forbes, *et al.*, 1990	Impacted third molar (1 or more) n = 269 Age 15+ years	RCT, double-blind, single then multiple oral dose, 6 parallel groups. General/local anaesthetic (unclear). Self-assessed at home at 0, 1 hours, then hourly for 6 hours. Returned 5 days later for review and debriefing.	PI (4-point scale) PR (5-point scale) Pain half gone? (y/n) Global rating (5-point scale) Time to remedication	Placebo (n = 34); paracetamol, 600 mg (n = 36).	All active medications significantly superior for all measures of total and peak analgesia.	Remedication allowed after 2 hours but asked to complete next evaluation (PI last or baseline score; PR = 0 used for all remaining time points). % remedicated by hour 6: placebo 33%; paracetamol 29%.	206 analysed. Exclusions; 3 lost to follow-up; 1 lost report card; 22 did not require medication; 8 remedicated despite having relief from study medication; 6 remedicated with only slight pain; 13 remedicated < 2 hours; 7 failed to follow instructions; 3 did not complete forms.	None serious reported. Numbers reporting adverse effects (number of effects): placebo 0/38; paracetamol 5/41 (5).	5
Honig & Murray, 1984	Postoperative (elective surgery, viz. abdominal, orthopaedic, rectal, thoracic, vascular) n = 116 Age range 19–87 years	RCT, single oral dose, 4 parallel groups. No information on anaesthesia, 4-hour washout prior to start. Interviewed in hospital by nurse observer at 0, 0.5, 1 hours then hourly for 6 hours.	PI (4-point scale) PR (5-point scale) 50% PR? (y/n) Global rating (5-point scale)	Placebo (n = 30); paracetamol, 600 mg (n = 28).	TOTPAR and global rating: paracetamol significantly superior to placebo ($p < 0.05$).	If remedicated, last score used for all further time points. Number of patients not remedicated in 6 hours: placebo 14/30; paracetamol 16/28.	No details given.	None severe except 1 severe dry mouth. Reported in all groups; primarily central nervous system and gastrointestinal effects.	3
Jain, *et al.*, 1986	Postoperative (general, gynaecological or orthopaedic surgery) n = 128 Age range 18–70 years	RCT, double-blind, single oral dose, 4 parallel groups. General anaesthetic. Trial drug given within 72 hours of surgery on request for analgesia. 4-hour washout prior to start. Assessed in hospital by nurse observer at 0, 15 and 30 minutes, then hourly for 6 hours.	PI (4-point scale) PR (5-point scale)	Placebo (n = 32); paracetamol, 650 mg (n = 30).	Paracetamol only significantly superior for maximum relief ($p < 0.05$).	Remedicated at < 2 hours, data excluded. If > 2 hours, last measurement used for all remaining time points. Total numbers remedicated: placebo 11/32; paracetamol 10/30; Nalbu 9/34; Combo 5/32.	122 analysed. Exclusions: 2 had improperly blinded drugs; 2 remedicated < 2 hours and 2 received interfering medication.	None serious reported. NSD between groups. Numbers reporting adverse effects (number of effects): placebo 6/32 (8) paracetamol 9/30 (9).	4

continued

TABLE 23 contd *Trials of paracetamol versus placebo*

Study	Condition and number of patients	Design, study duration and follow-up	Outcome measures	Dosing regimen	Analgesic outcome results	Remedication	Withdrawals and exclusions	Adverse effects	Quality score
Kiersch, *et al.*, 1994	Impacted third molar (3 or 4) n = 232 Age 14+ years	RCT, double-blind, single oral dose, 3 parallel groups. Local anaesthetic, 48-hour washout prior to start. Self-assessed in clinic for first 2 hours, then at home at 0, 20, 30, and 40 minutes, 1 hour then hourly for 12 hours.	PI (4-point scale) PR (5-point scale) Pain half gone? (y/n) Global rating (5-point scale) Time to remedication VAS PI (100 mm)	Placebo (n = 30); paracetamol, 1000 mg (n = 30).	Paracetamol significantly superior to placebo for most efficacy measures in first 6 hours.	"Patients were asked to allow 2 hours... before taking alternate medication." Time to remedication (median): placebo 2.0 hours; paracetamol 3.1 hours.	226 analysed. Exclusions: 1 experienced nausea and vomiting so did not ingest treatment; 2 did not require analgesia; 1 failed to follow instructions; 2 vomited within 10 minutes of taking trial drug.	None serious reported. Numbers reporting adverse effects (number of effects): placebo 13/45 (18); paracetamol 31/92 (35).	4
McQuay, *et al.*, 1988	Postoperative (elective orthopaedic surgery) n = 158 Age range 18–70 years	RCT, double-blind, single oral dose, 5 parallel groups. General anaesthetic. Trial drug given 1/2 days after surgery. 3-hour washout prior to start. Assessed in hospital by nurse observer at 0, 0.5, 1 and 1.5 hours then hourly for 6 hours.	PI (4-point scale, VAS and 8-word verbal rating) PR (5-point scale and VAS) Pain half gone? (y/n) Global rating (5-point scale, both patient and observer) Time to remedication Vital signs	Placebo (n = 30); paracetamol, 1000 mg (n = 30). NB: these are numbers after exclusions.	Paracetamol significantly superior to placebo for all integrated measures of efficacy.	If remedicated after 1 hour, PI scored at baseline and PR = 0. For patients. remedicated at < 6 hours: all active treatments significantly superior ($p < 0.01$) to placebo.	150 analysed. Exclusions: 2 discharged before end; 3 received drugs prohibited by protocol; 1 vomited intact medication within 15 minutes; 2 pain assessments inadequately completed.	None serious reported; NSD between groups. Numbers reporting adverse effects (number of effects): placebo 6/30 (8); paracetamol 6/30 (10).	4
Mehlisch, *et al.*, 1995	Third molar (at least 1 embedded) n = 240 Age 15+ years	RCT, double-blind, single oral dose, 3 parallel groups. Local anaesthetic, 12-hour washout prior to start. Self-assessed at 0, 15 and 45 minutes, 1 hour and 90 minutes, then hourly for 6 hours.	PI (4-point scale) PR (5-point scale) Global rating (5-point scale)	Placebo (n = 40); paracetamol, 1000 mg (n = 101).	Paracetamol significantly superior to placebo for all measures of efficacy.	If remedicated before 1 hour, data excluded from analysis. Value of 0 was assigned for PI and PR at all time points after remedication.% patients remedicating: placebo 88%; paracetamol 52%.	399 analysed. Exclusions: 1 failed to complete diary.	None serious reported. Numbers reporting adverse effects (number of effects): placebo 4/40 (?); paracetamol 17/101 (?).	3
Mehlisch & Frakes, 1984	Oral surgery (involving bone removal) n = 174 Age 16+ years	RCT, double-blind, single oral dose, 3 parallel groups. No information on anaesthesia except no long-acting i.m. or i.v. anaesthesia used. 4-hour washout period prior to start. Assessed in clinic by nurse observer at 30 minutes then hourly for 6 hours.	PI (4-point scale) PR (5-point scale) Pain half gone? (y/n) Global rating (5-point scale) Time to remedication	Placebo (n = 55); paracetamol, 1000 mg (n = 58).	Paracetamol significantly superior ($p < 0.05$) to placebo for all measures, and to aspirin for maximum PID ($p < 0.05$), maximum PR ($p < 0.03$) and global rating ($p < 0.02$).	Remedicated if required after 1 hour and patient considered treatment failure (no details on how data was handled). Full dichotomous data on times of medication in Table IV. Number of patients not remedicated before t > 6 hours: paracetamol 13/58; placebo 3/55.	162 analysed. Exclusions: 9 failed to comply with protocol; 3 lost to follow-up.	NSD in numbers for paracetamol and placebo – no other details given.	4
Schachtel, *et al.*, 1989	Episiotomy (post uncomplicated delivery) n = 115 Age range 16–37 years	RCT, double-blind, single oral dose, 3 parallel groups. 4-hour washout prior to start. Assessed (where and by whom not clear) at 0, 0.5, 1 hour, then hourly for 4 hours.	PI (4-point scale) PR (5-point scale) Global rating (5-point scale)	Placebo (n = 38); paracetamol, 1000 mg (n = 37).	Paracetamol significantly superior ($p < 0.05$) to placebo for TOTPAR, global rating and number of remedications.	If remedicated after 1 hour considered treatment failure; last/baseline PI and PR = 0 scored for remaining time points. Remedicated < 6 hours: paracetamol 13/37; placebo 22/38.	111 analysed. Exclusions: 4 remedicated but did not record at what time.	None reported. Placebo 0/37; paracetamol 0/37.	4
Sunshine, *et al.*, 1986	Impacted third molar n = 182 Age 16+ years	RCT, double-blind, single oral dose, 6 parallel groups. Local anaesthetic (lidocaine/epinephrine). 4-hour washout prior to start. Evaluations at 0, 0.5, 1, 2, 3 hours in clinic by single observer. At 4, 5, 6 hours self-assessed. 1 week later met with observer to review forms.	PI (4-point scale) PR (5-point scale) Global rating (4-point scale) Overall improvement (7-point scale) Time to remedication	Placebo (n = 30); paracetamol, 650 mg (n = 30).	Paracetamol significantly superior ($p < 0.05$) to placebo for: PID at 1–3 hours, SPID at 4 hours, PR at 2 hours and time to peak effect (see Table II).	If remedicated at < 1 hour, excluded. If at > 1 hour, last PI or baseline and PR = 0 used. Full dichotomous data in Table III. Number of patients remedicated: placebo 13/30; paracetamol 14/30.	182 analysed. No exclusions.	None serious reported; NSD between groups. Numbers reporting adverse effects: placebo 1/30; paracetamol 1/30.	5

continued

TABLE 23 contd *Trials of paracetamol versus placebo*

Study	Condition and number of patients	Design, study duration and follow-up	Outcome measures	Dosing regimen	Analgesic outcome results	Remedication	Withdrawals and exclusions	Adverse effects	Quality score
Sunshine, *et al.*, 1989	Episiotomy (multiparous in-patients) n = 200 Age 18+ years	RCT, double-blind, single oral dose, 3 parallel groups. Only included patients with severe pain. 4-hour washout prior to start. Evaluations in hospital by nurse-observer at 0, 0.5, 1 hours, then hourly for 6 hours (if asleep, woken). Interviews in patients' first language (Spanish).	PI (4-point scale) PR (5-point scale) Global rating (4-point scale) Overall improvement (7-point scale) Time to remedication	Placebo (n = 50); paracetamol, 650 mg (n = 75).	Paracetamol significantly superior ($p < 0.05$) to placebo for all measures of efficacy.	If remedicated at < 2 hours excluded. If remedicated at > 2 hours, PI last or baseline and PR = 0 used for remaining time points. No remedications at < 2 hours. Remedications at > 2 hours: placebo 8/50; paracetamol 2/75.	200 analysed. No exclusions.	Only 2 patients reported adverse efffect; they were in neither paracetamol nor placebo groups. Those reported were mild dizziness, sleepiness and sweating.	4
Sunshine, *et al.*, 1993	Postoperative (Caesarean section) n = 240 Age 18+ years	RCT, double-blind, single oral dose then multi-dose, 5 parallel groups. Only included patients with severe pain, 4-hour washout prior to start. Evaluation by same nurse-observer at 0, 0.5, 1 hours, then hourly for 8 hours. Interviews in patients' first language (Spanish).	PI (4-point scale) PR (5-point scale) Global rating (5-point scale) Time to meaningful relief	Placebo (n = 48); paracetamol, 650 mg (n = 48).	Paracetamol not significantly superior ($p < 0.05$) to placebo for any measure.	If remedicated at < 1 hour after first dose, dropped and replaced. If remedicated > 1 hour after first dose, eligible for repeat dose phase. Number of patients remedicated < 8 hours: placebo 35/48; paracetamol 42/48.	All 240 enrolled were analysed.	No details for single dose phase.	4
Winter, *et al.*, 1983	Oral surgery (various procedures) n = 168 Age range 16–75 years	RCT, double-blind, single oral dose, 4 parallel groups. 4-hour washout period prior to start. General and/or local anaesthetic. Self-assessed at 0, 0.5, 1, 2, 3 and 4 hours.	PI (4-point scale) PR (5-point scale) 50% PR? (y/n) Global rating (5-point scale)	Placebo (n = 41); paracetamol, 1000 mg (n = 41).	Paracetamol significantly superior to placebo for all measures of analgesic efficacy. Both produced significant analgesia as early as t = 0.5 hours.	Remedication allowed after 2 hours or if pain returned to pre-medication levels. 2 patients remedicated ≥ 2 hours (1 placebo and 1 paracetamol).	164 analysed. Exclusions: 3 protocol violations; 1 did not receive study medication.	None serious reported. Numbers reporting adverse effects: placebo 1/41 (severe headache); paracetamol 0/41.	4
Young, *et al.*, 1979	Postoperative (various elective procedures) n = 120 Age range 12–83 years	RCT, double-blind, single oral dose, 4 parallel groups. General anaesthetic (halothane/nitrous oxide/oxygen). 4-hour washout prior to start. Evaluated in hospital by single observer at 0, 0.5, 1, 2, 3 and 4 hours.	PI (4-point scale) PR (5-point scale) 2 Global ratings (5-point scale, both patient and observer opinions)	Study 1: Placebo (n = 29); paracetamol, 650 mg (n = 30).	Paracetamol only significantly superior to placebo ($p < 0.05$) at t = 2 hours PR score.	"Any concomitant or additional medication given was duly noted." No details of this data given or how it was handled.	119 analysed. Exclusion: 1 received analgesia within 2 hours of study.	None serious reported. Numbers reporting adverse effects: placebo 1/30 (sedation); paracetamol 3/30 (nausea).	4

Drug-related study withdrawals rarely occurred. One study[15] had three withdrawals, one patient on placebo and two on paracetamol. Another study[16] had two patients who withdrew on paracetamol. These studies reported a variable incidence of adverse events which were mild and transient, with no difference in incidence between paracetamol and placebo.

Paracetamol plus codeine versus placebo

A total of 20 reports of 20 trials were found which fulfilled the inclusion criteria; 721 patients were given paracetamol and 664 placebo. Details of the trials are given in *Table 25*. The doses administered were paracetamol, 300 mg, plus codeine, 30 mg, in five trials; paracetamol, 600 mg, plus codeine, 60 mg, in eight trials; paracetamol, 650 mg, plus codeine, 60 mg, in five trials, and paracetamol, 1000 mg, plus codeine, 60 mg, in two trials. One report on episiotomy provided dichotomous information on the overall patient global rating of pain relief.[17] This report was included. The proportion of patients with good or excellent pain relief was used. Pain relief was measured over 4–6 hours in 19 of the reports; one had observations for just 3 hours.

The variation in placebo response rate was from 0% to 72% of patients with at least 50% maxTOTPAR. The variation in response rate for all doses of paracetamol plus codeine was 20% to 83% (see *Figure 25,* page 73). The mean response rate for paracetamol, 600/650 mg, plus codeine, 60 mg, was 51% and for placebo, 21%.

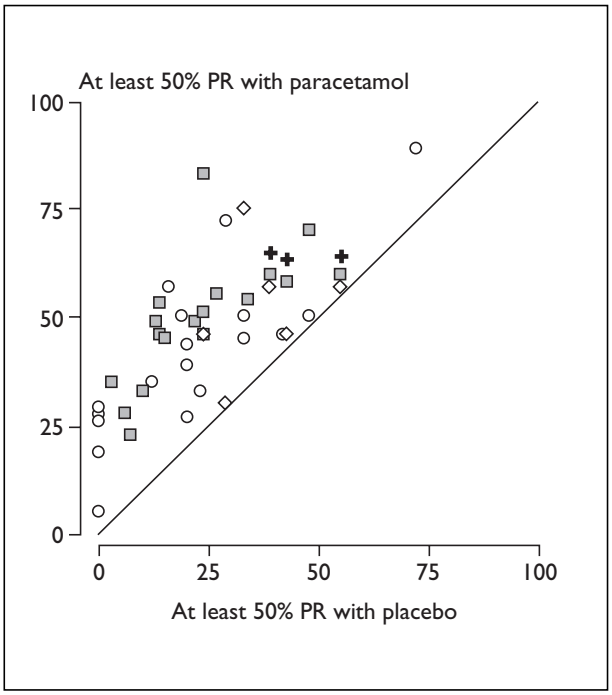


FIGURE 24 *Single dose studies of paracetamol against placebo (◇, 500 mg; ○, 600/650 mg; ▫, 1000 mg; +, 1500 mg)*

Doses of paracetamol were given first. Studies were predominantly in oral surgery (14 trials for paracetamol plus codeine against placebo and ten trials for paracetamol plus codeine against paracetamol alone).

Combining data across conditions, paracetamol, 300 mg, plus codeine, 30 mg, compared with placebo in single dose administration had an NNT for at least 50% pain relief of 5.3 (95% CI, 3.8–8.0), paracetamol, 600/650 mg, plus codeine, 60 mg, an NNT of 3.1 (95% CI, 2.6–3.8) and paracetamol, 1000 mg, plus codeine, 60 mg, an NNT of 1.9 (95% CI, 1.5–2.6), although in only two trials (see *Table 26,* page 74).

There were no serious adverse events which necessitated patient withdrawal from any study.

Paracetamol plus codeine versus paracetamol alone

A total of 12 reports of 12 trials were found which fulfilled the inclusion criteria; 395 patients were given paracetamol and 399 placebo. Four reports were identified which fulfilled the inclusion criteria except that they had pain outcomes other than five-point categorical pain relief scores. Details of the trials are given in *Table 27* (see page 75). Ten trials were in oral surgery and three in postsurgical pain. Doses were paracetamol, 600 mg, plus codeine, 60 mg, in seven trials, paracetamol, 650 mg, plus codeine, 60 mg, in four trials, and paracetamol, 1000 mg, plus codeine, 60 mg, in two trials. Pain relief was measured over 4–6 hours in 12 reports; one measured pain relief over 3 hours.

The variation in response rates for paracetamol alone was from 5% to 89% of patients with at least 50% maxTOTPAR. The variation in response rates to all doses of paracetamol plus codeine was 24–83%.

Only one of the reports had a lower 95% CI of the relative benefit that did not include 1 *(Figure 26,* see page 74). The combined relative benefit (fixed effects model) for this homogeneous data set was 1.25 (95% CI, 1.09– 1.43). Combining data across conditions, the NNT for addition of codeine, 60 mg, to all doses of paracetamol in single dose administration for at least 50% pain relief was 7.7 (95% CI, 5.1–17).

There were no serious adverse events which necessitated patient withdrawal from any study.

Comment

Paracetamol, 1000 mg, alone had an overall NNT of 4.6 for at least 50% pain relief compared with placebo in single dose administration. This means that one in every five patients with pain of moderate to severe intensity will get at least 50% pain relief; they would not have done had they been

TABLE 24 *Summary risk ratios and NNTs for trials of paracetamol against placebo*

Number of trials	Paracetamol dose (mg)	> 50% max TOTPAR on paracetamol	> 50% max TOTPAR on placebo	RB (95% CI)	NNT (95% CI)
6	500	194/353	109/296	1.4 (1.1–1.9)	5.6 (3.9–9.5)
17	600/650	243/594	125/573	1.7 (1.3–2.2)	5.3 (4.1–7.2)
20	1000	620/1376	207/907	2.3 (1.7–2.9)	4.6 (3.9–5.4)
3	1500	133/207	63/141	1.4 (1.2–1.9)	5.0 (3.3–11)

TABLE 25 *Studies of paracetamol plus codeine versus placebo*

Study	Condition and number of patients	Design, study duration and follow-up	Outcome measures	Dosing regimen	Analgesic outcome results	Remedication	Withdrawals and exclusions	Adverse effects	Quality score
Bentley & Head, 1987	Third molar bony impacted n = 128 Mean age: mid-20s	RCT, double-blind, single oral dose, 4 parallel groups. No information on anaesthesia except 'no sedative or narcotic agents were used before, during or after surgery'. Self-assessed at home at 0, 1, 2, 3, 4, 5 hours; reports posted to investigator.	PI (10-point scale) PR (5-point scale)	Placebo (n = 17); paracetamol, 1000 mg, + codeine, 60 mg, (n = 41)	Combination significantly superior to placebo for all measures of efficacy, but not significantly different from paracetamol for any measure.	Rescue analgesic: Tylenol no. 3. Patients who remedicated < 5 hours: last PI and PR scores carried on for all further time points.	120 analysed. Exclusions: 3 did not take medication; 1 took only a portion; 1 took no medication until day after surgery; 1 remedicated after 30 minutes; 1 vomited within 30 minutes of taking medication; 1 patient did not return forms.	53 had 1 or more; 86 reported in total, majority being dizziness, drowsiness, nausea and vomiting. No significant difference between treatment groups. Numbers reporting adverse effects (number of effects): placebo 9/19 (16); paracetamol + codeine 15/42 (24).	3
Cooper & Beaver, 1976	Impacted third molar n = 216 Age: 16+ years	RCT, double-blind, 2 single oral dose studies (second is paracetamol), 4 parallel groups. No information on anaesthesia. Self-assessed at home at 0, 1, 2, 3 hours (questionnaire posted to investigator).	PI (4-point scale) PR (5-point scale) Pain half gone? (y/n) Had the patient fallen asleep during the hour? (y/n)	Study 2: Placebo (n = 40); paracetamol, 600 mg, + codeine, 60 mg (n = 40)	Parametric analysis concluded paracetamol significantly superior to placebo ($p < 0.01$) and non-parametric factorial analysis showed codeine had effect for hour 1 ($p < 0.01$) and peak PI ($p < 0.05$). NSD between combination and its constituents	Remedication consisted of taking second envelope and evaluating it in same way as first; unclear how data was then handled. 112/160 patients required remedication.	160 analysed. Exclusions: 30 patients did not return forms, 12 filled in the forms incorrectly, 4 took concomitant medication, 6 required no medication and 4 were randomly deleted to even out number of patients on each treatment.	None serious reported. Most common were drowsiness, nausea and headache. Numbers reporting adverse effects (number of effects): placebo 5/40 (5); paracetamol + codeine 11/40 (11).	5
Cooper, *et al.*, 1981	Impacted third molar n = 248 Mean age: early 20s	RCT, double-blind, single oral dose, 5 parallel groups. General or local anaesthetic. Self-assessed at home at 0, 1, 2, 3, 4 hours (questionnaire).	PI (4-point scale) PR (5-point scale) Global rating (5-point scale) 50% relief of baseline pain? (y/n) Time to remedication	Placebo (n = 37); paracetamol, 650 mg, + codeine, 60 mg, (n = 42)	All active treatments significantly superior to placebo for all measures. Combination slightly more effective than paracetamol alone but difference not significant.	Remedication after first hour if needed (last score used for all further time points). Remedication < 4 hours: placebo 20/37; paracetamol 2/37.	200 analysed. Exclusions: 17 did not ingest medication, 31 ingested medication but violated protocol (remedication before hour 1 observations; constant deviation of more than 15 minutes from evaluation times, not returning questionnaire and lost to follow-up).	None serious reported. Numbers reporting adverse effects (number of effects): placebo 4/37 (5); paracetamol + codeine 10/42 (10).	4
Cooper, *et al.*, 1988	Impacted third molar n = 165 Age range: 18—57 years	RCT, double-blind, single oral dose, 3 parallel groups, single centre and 1 surgeon. Local anaesthetic with sedation and/or nitrous oxide. 4-hour washout prior to start. Self-assessed at home at 0, 0.5, 1, 2, 3, 4, 5, 6 hours (diary).	PI (4-point scale) PR (5-point scale) Pain half gone? (y/n) Global rating (5-point scale) Time to remedication	Placebo (n = 40); paracetamol, 600 mg, + codeine, 60 mg (n = 31)	Combination significantly superior to placebo for every measure and to paracetamol for TOTPAR ($p < 0.05$).	Remedication after first hour (last or baseline score used for all further time points).	143 analysed. Exclusions: 11 lost to follow-up, 8 did not require medication and 3 for 'various protocol violations'.	None serious reported. Number of effects reported: placebo 3; paracetamol + codeine 4.	4
Cooper & Kupperman, 1991	Removal of one or more impacted teeth n = 247 Age: 'young adults'	RCT, double-blind, single oral dose, 6 parallel groups. Local anaesthetic (lidocaine + epinephrine) with i.v. diazepam and methohexital – on occasion nitrous oxide was also used. Self-assessed at home at 0, 0.5 hours then hourly for 6 hours (diaries).	PI (4-point scale) PR (5-point scale) Global rating (5-point scale) Time to remedication	Placebo (n = 44); paracetamol, 650 mg, + codeine, 60 mg (n = 39)	Combination significantly superior to placebo for most measures and to paracetamol for TOTPAR and global rating.	t > 1 hour before remedication; data included and baseline or last score (most severe) used for all further time points.	226 analysed. Exclusions; 13 did not require medication, 3 lost to follow-up, 2 remedication. with slight pain before hour 2 observations, 2 remedications. Before first hour observations, 1 fell asleep for over 2 hours.	All mild. Numbers reporting adverse effects (number of effects): placebo 7/44 (9); paracetamol + codeine 8/39 (11).	3

continued

TABLE 25 contd *Studies of paracetamol plus codeine versus placebo*

Study	Condition and number of patients	Design, study duration and follow-up	Outcome measures	Dosing regimen	Analgesic outcome results	Remedication	Withdrawals and exclusions	Adverse effects	Quality score
Desjardins, *et al.*, 1986	Oral surgery n = 137 Age: 18+ years	RCT, double-blind, single oral dose, 3 parallel groups, local anaesthetic. 4-hour washout prior to start. Self-assessed at home at 0, 0.5 hours, then hourly for 6 hours.	PI (4-point scale) PR (5-point scale) 50% PR (y/n) Global rating (5-point scale) Anxiety (4-point scale) Relaxation (4-point scale)	Placebo (n = 41); paracetamol, 300 mg, + codeine, 30 mg (n = 39)	Combination only significantly superior to placebo for global rating and total anxiety ($p < 0.05$).	If t ≥ 1 hour before remedication, data included and last score for PI and PR used for all further time points.	123 analysed. Exclusions:14 did not medicate, lost to follow-up, or provided uninterpretable results.	None serious reported. Numbers reporting adverse effects (number of effects): placebo 4/41 (4); paracetamol + codeine 2/39 (3).	4
Dionne, *et al.*, 1994	Impacted third molar n = 135 Age: 16+ years	RCT, double-blind, single oral dose, 5 parallel groups. General and local anaesthetic. 4-hour washout prior to start. Self-assessed at clinic for at least the first 2 hours, then at home hourly for 6 hours.	PI (4-point scale) PR (5-point scale) Global rating (5-point scale) Time to remedication	Placebo (n = 25); paracetamol, 650 mg, + codeine, 60 mg (n = 24)	Neither paracetamol nor combination significantly different from placebo for any measure of analgesia.	t ≥ 2 hours before remedication; data included and baseline used for all further time points.	124 analysed. Exclusions;4 previously enrolled in the study, 3 remedicated before t = 2 hours, 2 lost to follow up, 1 ineligible because of codeine sensitivity.	All mild. Numbers reporting adverse effects (number of effects): placebo 5/25 (5); paracetamol + codeine 9/24 (10).	3
Forbes, *et al.*, 1982	Impacted third molar (1 or more) n = 177 Age: 15+ years	Double-blind, single oral dose, 5 parallel groups. No information on anaesthesia. Self-assessed at home at 0, 1 hours, then hourly for 12 hours.	PI (4-point scale) PR (5-point scale) 50% PR (y/n) Global rating (5-point scale) Time to remedication	Placebo (n = 30); paracetamol, 600 mg, + codeine, 60 mg (n = 31)	Combination significantly superior to placebo for all measures.	Excluded if remedicated at < 2 hours. If > 2 hours but < 12 hours, PR = 0 and PI baseline or last. Patients remedicatedat > 2 hours but < 12 hours: placebo 93%; paracetamol 85%.	159 analysed. Exclusions: 4 lost to follow-up, 3 did not take medication, 7 took remedication < 2 hours, 1 took medication with mild pain and 3 did not complete the forms properly.	No active treatment produced more than placebo. None were serious.	3
Forbes, *et al.*, 1983	Postoperative (general, gynaecological or orthopaedic surgery) n = 132 Age: 18+ years	RCT, double-blind, single oral dose, 5 parallel groups. General anaesthetic, given on request first day able to take oral analgesia. Self-assessed in hospital at 0, 0.5, 1, 1.5, 2 hours, then hourly for 12 hours. Single nurse observer present for first 6 hours, next 6 hours monitored by ward staff.	PI (4-point scale) PR (5-point scale) Pain half gone? (y/n) Global rating (5-point scale) Time to remedication	Placebo (n = 26); paracetamol, 600 mg, + placebo, 60 mg (n = 26)	For 6-hour data, combination significantly superior to placebo for all measures; for 12-hour data, all except SPID.	Patients remedicated on demand at < 2 hours excluded. If remedication < 12 hours, PR = 0, PI = baseline or last score for remaining time points.	132 analysed. There were no exclusions.	None serious reported. Most common were drowsiness, dizziness and dry mouth. Numbers reporting adverse effects (number of effects): placebo 4/26 (5); paracetamol + codeine 11/26 (16).	5
Forbes, *et al.*, 1986	Impacted third molar (1 or more) n = 146 Age: 15+ years	RCT, double-blind, single oral dose, 3 parallel groups, local anaesthetic. Self-assessed at home at 0, 1 hours, then hourly for 6 hours (diary); returned 5 days later for review and debriefing.	PI (4-point scale) PR (5-point scale) 50% PR (y/n) Global rating (5-point scale) Anxiety (4-point scale) Relaxation (4-point scale)	Placebo (n = 38); paracetamol, 300 mg, + placebo, 30 mg (n = 43)	Combination significantly superior to placebo for all measures of efficacy.	If remedicated at < 2 hours, excluded. If > 2 hours but < 6 hours, PR = 0 and PI baseline or last (which ever was greater).	122 analysed. Exclusions:1 did not return the form, 6 did not need analgesia and 17 had invalid data.	None serious reported. Numbers reporting adverse effects (number of effects): placebo 9/46 (11); paracetamol + codeine 6/46 (8).	5
Forbes, *et al.*, 1989	Impacted third molar (1 or more) n = 107 Age: 15+ years	RCT, double-blind, single oral dose, 4 parallel groups at 2 centres. General and local anaesthetic (unclear). Self-assessed at home at 0, 1, 2 hours, then hourly for 12 hours or until remedication (diary); returned 5 days later for review and debriefing.	PI (4-point scale) PR (5-point scale) Global rating (5-point scale) Time to remedication	Placebo (n = 23); paracetamol, 600 mg, + codeine, 60 mg (n = 17)	At 4 hours: combination significantly superior to placebo paracetamol for TOTPAR ($p < 0.05$).	Patients could remedicate after 2 hours but were asked to complete next evaluation before doing so (PI last or baseline score, PR = 0 for all further time points). Patients remedicating by hour 12: paracetamol 95%; placebo 91%.	88 analysed. Exclusions: 9 did not take medication, 2 remedicated before the 2-hour point, 1 remedicated with slight pain, 4 did not complete their evaluation, 1 took only part of the medication and 2 remedicated despite getting some relief from study medication.	None serious reported. Numbers reporting adverse effects (number of effects): placebo 2/26 (2); paracetamol + codeine 1/17 (1). NB: includes those reported post-remedication.	5

continued

TABLE 25 contd *Studies of paracetamol plus codeine versus placebo*

Study	Condition and number of patients	Design, study duration and follow-up	Outcome measures	Dosing regimen	Analgesic outcome results	Remedication	Withdrawals and exclusions	Adverse effects	Quality score
Forbes, *et al.*, 1990a	Impacted third molar (1 or more) n = 162 Age: 15+ years	RCT, double-blind, single oral dose, 4 parallel groups, local anaesthetic. Self-assessed at home at 0, 1 hours, then hourly for 6 hours (diary); returned 5 days later for review and debriefing.	PI (4-point scale) PR (5-point scale) Pain half gone? (y/n) Global rating (5-point scale) Time to remedication	Placebo (n = 32); paracetamol, 600 mg, + codeine, 60 mg (n = 27)	Combination significantly superior to placebo for all measures of efficacy.	If remedication < 2 hours, excluded. If > 2 hours but < 6 hours, PR = 0 and PI baseline or last which ever was greater. Number remedicating before t = 6 hours: placebo 27; combination 23.	128 analysed. Exclusions: 1 failed to return form, 19 did not require analgesia and 14 had invalid data.	None serious reported. Numbers reporting adverse effects (number of effects): placebo 5/34 (6); paracetamol + codeine 9/31 (12).	5
Forbes, *et al.*, 1990b	Impacted third molar (1 or more) n = 269 Age: 15+ years	RCT, double-blind, single then multiple oral dose, 6 parallel groups. General/local anaesthetic (unclear). Self-assessed at home at 0, 1 hours, then hourly for 6 hours; returned 5 days later for review and debriefing.	PI (4-point scale) PR (5-point scale) Pain half gone? (y/n) Global rating (5-point scale) Time to remedication	Placebo (n = 34); paracetamol, 600 mg, + codeine, 60 mg (n = 38)	All active medications significantly superior for all measures of total and peak analgesia.	Remedication allowed after 2 hours but asked to complete next evaluation (PI last or baseline score, PR = 0, used for all remaining time points). Patients remedicating. by hour 6: placebo 33%; paracetamol 29%.	206 analysed. Exclusions; 3 lost to follow-up, 1 lost report card, 22 did not need medication. 8 remedicated despite having relief from study mediation, 6 remedicated with only slight pain, 13 remedicated at < 2 hours, 7 failed to follow instructions, 3 did not complete forms.	None serious reported. Numbers reporting adverse effects (number of effects): placebo 0/38; paracetamol + codeine 8/40 (9).	5
Forbes, *et al.*, 1994	Impacted third molar (1 or more) n = 324 Age: 15+ years	RCT, double-blind, single oral dose, 3 parallel groups, local anaesthetic. Patients self-assessed at home at 0, 0.5, 1 hours, then hourly for 6 hours (diary); returned 5 days later for review and debriefing.	PI (4-point scale) PR (5-point scale) Pain half gone? (y/n) Global rating (5-point scale) Time to remedication	Placebo (n = 45); paracetamol, 300 mg, + placebo, 30 mg (n = 93)	Combination significantly superior to placebo for all measures of mean and peak analgesia.	If remedication at < 2 hours, excluded. If > 2 hours but < 6 hours, PR = 0 and PI baseline or last, whichever was greater.	232 analysed. Exclusions: 1 lost to follow-up, 32 did not require analgesia and 51 had invalid data.	None serious reported. Numbers reporting adverse effects (number of effects): placebo 10/65 (12); paracetamol + codeine 18/107 (21).	5
Heidrich, *et al.*, 1985	Orthopaedic surgery n = 120 Age range: 18–65 years	RCT, double-blind, single oral dose, 3 parallel groups. No information on anaesthesia, 3-hour washout prior to start. Interviewed in hospital by nurse observer at 0, 0.5, 1 hours, then hourly for 6 hours.	PI (4-point scale) PR (5-point scale) Global rating (5-point scale) VAS PI and PR McGill questionnaire Mood questionnaire and VAS	Placebo (n = 40); paracetamol, 300 mg, + placebo, 30 mg, (n = 40)	Orthogonal analyses of variance showed combination gave greater relief from pain than placebo.	No information on remedication.	No information on withdrawals or exclusions.	No differences between treatments in terms of side-effects. No patient withdrew because of adverse events.	2
Honig & Murray, 1984	Postoperative (elective surgery: abdominal, orthopaedic, rectal, thoracic, vascular) n = 116 Age range: 19–87 years	RCT, double-blind, single oral dose, 4 parallel groups. No information on anaesthesia, 4-hour washout prior to start. Interviewed in hospital by nurse observer at 0, 0.5, 1 hours, then hourly for 6 hours.	PI (4-point scale) PR (5-point scale) 50% PR (y/n) Global rating (5-point scale)	Placebo (n = 30); paracetamol, 600 mg, + codeine, 60 mg (n = 30)	Combination significantly superior to placebo for most measures of efficacy.	If remedicated, last score used for all further time points. Number who did not remedicate in 6 hours: placebo 14/30; paracetamol 16/28.	No details given.	None severe except 1 (severe dry mouth). Primarily central nervous system and gastrointestinal effects in all groups.	3
Petti, 1985	Orthopaedic or general surgery n = 141 Age range: 18–80 years	RCT, double-blind, single oral dose, 4 parallel groups. No information on anaesthesia, 4-hour washout prior to start. Interviewed in hospital by observer at 0, 0.5, 1 hours, then hourly for 6 hours. NB: all patients had a baseline PI of 2 (moderate).	PI (4-point scale) PR (5-point scale) Global rating (5-point scale) Severity of adverse effects (5-point scale)	Placebo (n = 32); paracetamol, 300 mg, + placebo, 30 mg (n = 31)	Does not say anything directly about combination.	Remedication allowed after 2 hours. If remedicated, scores of PR = 0 and PI = 2 allocated and patient excluded from further evaluations.	129 analysed. Exclusions: 12 excluded for protocol violations.	None serious reported. Only 1 reported (dry mouth) in paracetamol + codeine group.	2

continued

TABLE 25 contd *Studies of paracetamol plus codeine versus placebo*

Study	Condition and number of patients	Design, study duration and follow-up	Outcome measures	Dosing regimen	Analgesic outcome results	Remedication	Withdrawals and exclusions	Adverse effects	Quality score
Sunshine, *et al.*, 1986	Impacted third molar n = 182 Age: 16+ years	RCT, double-blind, single oral dose, 6 parallel groups. Local anaesthetic (lidocaine/ epinephrine); 4-hour washout prior to start. Evaluations at 0, 0.5, 1, 2, and 3 hours in clinic by single observer. Self-assessed at 4, 5 and 6 hours. Met with observer 1 week later and reviewed forms.	PI (4-point scale) PR (5-point scale) Global rating (4-point scale) Overall improvement (7-point scale) Time to remedication	Placebo (n = 30); paracetamol, 650 mg, + codeine, 60 mg (n = 31)	Combination significantly superior to placebo for most measures of efficacy ($p < 0.05$).	If at < 1 hour excluded. If > 1 hour, last PI or baseline and PR = 0 used (Full dichotomous data: Table III. Number remedicated: placebo 13/30; paracetamol 14/30.	182 analysed. No exclusions.	None serious reported; NSD between groups. Numbers reporting adverse effects: placebo 1/30; paracetamol + codeine 3/31.	5
Turek & Baird, 1988	Elective surgery n = 161 Age: 18+ years	RCT, double-blind, single oral dose, 4 parallel groups. No information on anaesthesia, 3 hour washout prior to start. Interviewed in hospital by nurse observer at 0, 0.5, 1 hours then hourly for 6 hours.	PI (4-point scale) PR (5-point scale) Global rating (5-point scale) Subjective assessment of improvement (7-point scale)	Placebo (n = 41); paracetamol, 650 mg, + codeine, 60 mg (n = 39)	Combination significantly superior to placebo for most measures of efficacy.	If remedicated at < 1 hours, excluded. If > 1 hours but < 6 hours, PR = 0 and PI = baseline allocated for all further time points.	160 analysed. Exclusions: 1 from placebo for taking concomitant medication.	None serious reported. Numbers reporting adverse effects (number of effects): placebo 4/41 (6); paracetamol + codeine 11/39 (20).	3

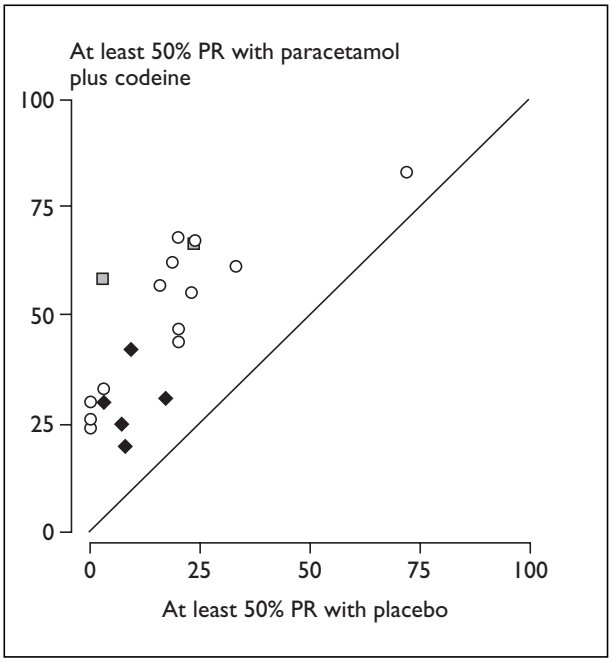


FIGURE 25 *Single dose studies of paracetamol plus codeine against placebo (◆, 300 mg + 60 mg; ○, 600/650 mg + 60 mg; □, 1000 mg + 60 mg)*

given placebo. The equivalent NNT at 600/650 mg was 5.3, indicating lower efficacy, although the dose response was not significant.

Paracetamol, 600/650 mg, plus codeine, 60 mg, compared with placebo in single dose administration had an NNT of 3.1 for at least 50% pain relief, meaning that one in every three patients with pain of moderate to severe intensity will get at least 50% pain relief; they would not have done so had they been given placebo *(Table 26)*. There was no overlap between the 95% CI of the NNT for paracetamol, 600/650 mg, plus codeine, 60 mg, in 13 trials (95% CI, 2.6–3.9) and that of paracetamol, 600/650 mg, alone in 17 trials (95% CI, 4.1–7.2). This indicates that addition of codeine, 60 mg, provides a substantial increase in analgesia in single dose administration. This is demonstrated clearly in *Figure 27* (see page 77), where wide CIs accompany point estimates of the NNT in trials with small numbers of patients. Despite this, paracetamol combined with codeine, 60 mg, is clearly a powerful analgesic.

The extra analgesic effect of adding codeine, 60 mg, to paracetamol can be estimated in two ways. Since both paracetamol alone and paracetamol plus codeine were compared with placebo, then any increased response rate (proportion of patients with at least 50% pain relief) may be ascribed to the addition of codeine. For paracetamol, 600/650 mg, alone against placebo, the difference between active (42%) and control (23%) response rates was 19%. For paracetamol, 600/650 mg, plus codeine, 60 mg, the difference between active (51%) and control (21%) response rates was 33%. Thus the extra 11% response was due to the addition of codeine, 60 mg.

There were also direct comparisons of paracetamol (all doses) plus codeine, 60 mg, with the same dose of paracetamol alone. Here again, any increased response rate can be ascribed to the addition of codeine. For paracetamol plus codeine, 60 mg, versus the same dose of paracetamol the difference between active (55%) and control (41%) response rates was 14%. This agreement between direct and indirect measures helps to justify the

TABLE 26 *Summary risk ratios and NNTs for trials of paracetamol and codeine against placebo and paracetamol alone*

Number of trials	Drug dose (mg) paracetamol +codeine	> 50% maxTOTPAR on paracetamol + codeine	> 50% maxTOTPAR on placebo	> 50% maxTOTPAR on paracetamol alone	RB (95% CI)	NNT (95% CI)
Paracetamol + codeine versus placebo						
5	300 + 30	69/246	17/196		3.0 (1.8–5.0)	5.3 (3.8–8.0)
13	600/650 + 60	200/398	80/418		2.6 (1.7–4.0)	3.1 (2.6–3.9)
2	1000 + 60	48/77	5/50		6.2 (0.8–47)	1.9 (1.5–2.6)
Paracetamol + codeine versus same dose of paracetamol alone						
10	600/650 + 60	165/309		129/313	1.3 (1.0–1.6)	8.3 (5.0–23)
2	1000 + 60	57/86		44/86	1.3 (1.1–1.5)	6.7 (3.4–174)
12	All doses + 60	222/395		173/399	1.2 (1.1–1.5)	7.7 (5.1–17)

meta-analytical methods. The extra 14% response for codeine, 60 mg, corresponds to an NNT for at least 50% pain relief in single dose administration of 7.7 (95% CI, 5.1–17) *(Table 26)*. This means that for every eight patients given paracetamol, 600/650 mg, plus codeine, 60 mg, one extra will achieve at least 50% pain relief who would not have done had they received paracetamol, 600/650 mg, alone.

The variation in placebo and active response rates was large, but this degree of variation is common in pain studies,[18] as well as in studies with more objective outcomes like postoperative vomiting,[19] and in the response of infants to pulmonary surfactant.[20] The variability in both the placebo and active response rates (see *Figures 24* and *25)* underpins the use of standard methods in pain research, where sensitivity of the model is demonstrated by separation of standard analgesic from placebo. This variability also emphasises both the need to include placebo groups in analgesic trials, and the need to understand better those factors that contribute to the variability in placebo responses in pain.

The power of the systematic review method is demonstrated here in several ways. The analgesic effect of paracetamol at two doses has been determined with confidence from all the available published data. The rather slight effects of codeine added to paracetamol (which are difficult to measure in single trials with limited numbers of patients) has been confirmed in direct and indirect comparison.

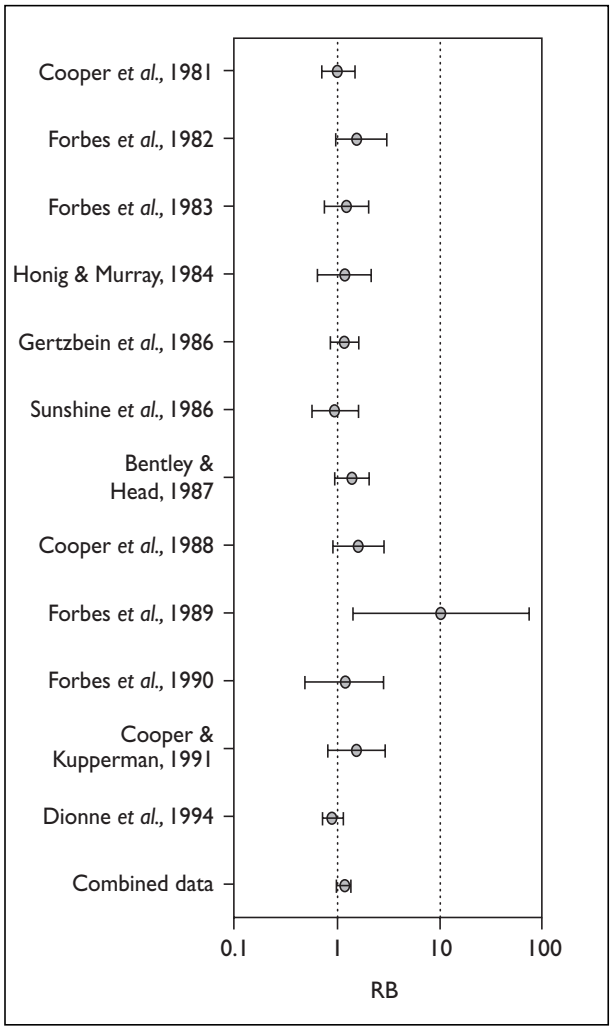


FIGURE 26 *Relative benefit for individual trials of codeine, 60 mg, added to paracetamol, compared with the same dose of paracetamol alone*

TABLE 27 *Studies of paracetamol plus codeine versus paracetamol*

Study	Condition and number of patients	Design, study duration and follow-up	Outcome measures	Dosing regimen	Analgesic outcome results	Remedication	Withdrawals and exclusions	Adverse effects	Quality score
Bentley & Head, 1987	Third molar bony impacted n = 128 Mean age: mid-20s	RCT, double-blind, single oral dose, 4 parallel groups. No information on anaesthesiaexcept 'no sedative or narcotic agents were used before, during or after surgery'. Self-assessed at home at 0, 1, 2, 3, 4, 5 hours then reports posted to investigator.	PI (10-point scale) PR (5-point scale)	Paracetamol, 1000 mg, + codeine, 60 mg (n = 41); paracetamol, 1000 mg (n = 41).	Both paracetamol and combination significantly superior to placebo for all measures of efficacy; combination not significantly different from paracetamol for any measure.	Rescue analgesic: Tylenol no. 3. For patients who remedicated at < 5 hours, last PI and PR scores carried on for all further time points.	120 analysed. Exclusions: 3 did not take medication, 1 took only a portion, 1 took no medication until day after surgery, 1 remedicated after 30 minutes, 1 vomited within 30 minutes of taking medication and 1 did not return forms.	53 had one or more, 86 reported in total; majority were dizziness, drowsiness, nausea and vomiting. NSD between treatment groups. Numbers reporting adverse effects (number of effects): paracetamol + codeine 15/42 (24); paracetamol 21/42 (31).	3
Cooper & Beaver, 1976	Impacted third molar n = 216 Age: 16+ years	RCT, double-blind, 2 single oral dose studies (study 2 is paracetamol), 4 parallel groups. No information on anaesthesia. Self-assessed at home at 0, 1, 2, 3 hours (questionnaire) then posted to investigator.	PI (4-point scale) PR (5-point scale) Pain half gone? (y/n) Had patient fallen asleep during the hour? (y/n)	Study 2: paracetamol, 600 mg, + codeine, 60 mg, (n = 40); paracetamol, 600 mg (n = 40).	Parametric analysis concluded paracetamol significantly superior to placebo ($p < 0.01$); non-parametric factorial analysis showed codeine had effects for first hour ($p < 0.01$) and peak PID ($p < 0.05$). No significant difference found between combination and its constituents.	Remedication consisted of taking second envelope and evaluating it in same way as the first; unclear how data were then handled. 112/160 patients required remedication.	160 analysed. Exclusions: 30 did not return forms, 12 filled forms in improperly, 4 took concomitant medication, 6 required no medication, 4 were randomly deleted to even out numbers of patients on each treatment.	None serious reported. Most common were drowsiness, nausea and headache. Numbers reporting adverse effects (number of effects): paracetamol + codeine 11/40 (11); paracetamol 5/40 (7).	5
Cooper, *et al.*, 1981	Impacted third molar n = 248 Mean age: early 20s	RCT, double-blind, single oral dose, 5 parallel groups. General/local anaesthesia. Self-assessed at home at 0, 1, 2, 3, 4 hours (questionnaire).	PI (4-point scale) PR (5-point scale) Global rating (5-point scale) 50% relief of baseline pain? (y/n) Time to remedication	Paracetamol, 650 mg, + codeine, 60 mg (n = 42); paracetamol, 650 mg (n = 37).	All active treatments significantly superior to placebo for all measures. Combination was slightly more effective than paracetamol alone but difference was not significant.	Remedicated after first hour if needed (last score used for all further time points). Remedication < 4 hours: placebo 20/37; paracetamol 2/37.	200 analysed. Exclusions: 17 did not ingest medication, 31 ingested medication but violated protocol (remedicated before first hour observation, constant deviation of more than 15 minutes from evaluation times, not returning questionnaire and lost to follow-up).	None serious reported. Numbers reporting adverse effects (number of effects): paracetamol + codeine 10/42 (10); paracetamol 12/37 (15).	4
Cooper, *et al.*, 1988	Impacted third molar n = 165 Age range: 18–57 years	RCT, double-blind, single oral dose, 3 parallel groups, single centre and 1 surgeon. Local anaesthetic with sedative and/or nitrous oxide. 4-hour washout prior to start. Self-assessed at home at 0, 0.5, 1, 2, 3, 4, 5, 6 hours (diary).	PI (4-point scale) PR (5-point scale) Pain half gone? (y/n) Global rating (5-point scale) Time to remedication	Paracetamol, 600 mg, + codeine, 60 mg (n = 31); paracetamol, 600 mg (n = 36).	Combination significantly superior to placebo for every measure and to paracetamol for TOTPAR. Paracetamol appeared clinically more effective than placebo but not significant. ($p < 0.05$).	Remedicated after first hour (last or baseline score used for all further time points).	143 analysed. Exclusions: 11 lost to follow-up, 8 did not require medication, 3 for 'various protocol violations'.	None serious reported. Numbers reporting adverse effects: paracetamol + codeine 4; paracetamol 8.	4
Cooper & Kupperman, 1991	Removal of one or more impacted teeth n = 247 Age: 'young adults'	RCT, double-blind, single oral dose, 6 parallel groups. Local anaesthetic (lidocaine + epinephrine) with i.v. diazepam and methohexital (nitrous oxide also used on occasion). Self-assessed at home at 0, 0.5 hours then hourly for 6 hours (diary).	PI (4-point scale) PR (5-point scale) Global rating (5-point scale) Time to remedication	Paracetamol, 650 mg, + codeine, 60 mg (n = 39); paracetamol, 650 mg (n = 37).	Paracetamol was only active drug not significantly superior to placebo for any measure. Combination significantly superior to placebo for most measures and to paracetamol for TOTPAR and global rating.	t > 1 hour before remedication; data included and baseline or last score (most severe) used for all further time points.	226 analysed. Exclusions; 13 did not require medication, 3 lost to follow-up, 2 remedicated with slight pain before hour 1 observations, 2 remedicated before hour 1 observations, 1 fell asleep for over 2 hours.	All reported mild. Numbers reporting adverse effects (number of effects): paracetamol + codeine 8/39 (11); paracetamol 6/37 (7).	3

continued

TABLE 27 contd *Studies of paracetamol plus codeine versus paracetamol*

Study	Condition and number of patients	Design, study duration and follow-up	Outcome measures	Dosing regimen	Analgesic outcome results	Remedication	Withdrawals and exclusions	Adverse effects	Quality score
Dionne, *et al.*, 1994	Impacted third molar n = 135 Age: 16+ years	RCT, double-blind, single oral dose, 5 parallel groups. General/local anaesthesia. 4-hour washout prior to start. Self-assessed at clinic for at least first 2 hours then at home hourly for 6 hours.	PI (4-point scale) PR (5-point scale) Global rating (5-point scale) Time to remedication	Paracetamol, 650 mg, + codeine, 60 mg (n = 24); paracetamol, 650 mg (n = 27).	Neither paracetamol nor combination were significantly different from placebo for any measure of analgesia.	t ≥ 2 hours before remedicating; data included and baseline used for all further time points.	124 analysed. Exclusions: 4 previously enrolled in study, 3 remedicated before t = 2 hours, 2 lost to follow-up, 1 ineligible because of codeine sensitivity.	All reported mild. Numbers reporting adverse effects (number of effects): paracetamol + codeine 9/24 (10); paracetamol 7/27 (9).	3
Forbes, *et al.*, 1982	Impacted third molar (1 or more) n = 177 Age: 15+ years	Double-blind, single oral dose, 5 parallel groups. No information on anaesthesia. Self-assessed at home at 0, 1 hours then hourly for 12 hours.	PI (4-point scale) PR (5-point scale) 50% PR? (y/n) Global rating (5-point scale) Time to remedication	Paracetamol, 600 mg, + codeine, 60 mg (n = 31); paracetamol, 600 mg (n = 34).	At 4 hours paracetamol significantly superior to placebo for PR, peak PR and 50% PR ($p < 0.01$). Combination significantly superior to placebo for all measures.	If remedicated < 2 hours, excluded. If > 2 < 12 hours, PR = 0 and PI baseline or last. % patients remedicated at > 2 < 12 hours: placebo 93%; paracetamol 85%.	159 analysed. Exclusions: 4 lost to follow-up, 3 did not take medication, 7 took remedication at < 2 hours, 1 medicated with mild pain, 3 did completed forms improperly.	None of active treatments produced more than placebo. None were serious.	3
Forbes, *et al.*, 1983	Postoperative (general, gynaecological or orthopaedic surgery) n = 132 Age: 18+ years	RCT, double-blind, single oral dose, 5 parallel groups. General anaesthetic, given on request on first day able to take oral analgesia. Self-assessed in hospital at 0, 0.5, 1, 1.5, 2 hours then hourly for 12 hours. Single nurse observer present for first 6 hours, next 6 hours monitored by ward staff.	PI (4-point scale) PR (5-point scale) Pain half gone? (y/n) Global rating (5-point scale) Time to remedication	Paracetamol, 600 mg, + placebo, 60 mg, (n = 26); paracetamol, 600 mg (n = 26).	For 6-hour data both paracetamol and combination significantly superior to placebo for all measures and for 12-hour data, for all except SPID.	Patients remedicated on demand but if at < 2 hours, excluded. If remedicated < 12 hours, PR = 0, and PI = baseline or last score for remaining time points.	132 analysed; no exclusions.	None serious reported. Most common were drowsiness, dizziness and dry mouth. Numbers reporting adverse effects (number of effects): paracetamol + codeine 11/26 (16); paracetamol 11/26 (11).	5
Forbes, *et al.*, 1989	Impacted third molar (1 or more) n = 107 Age: 15+ years	RCT, double-blind, single oral dose, 4 parallel groups at 2 centres. General/local anaesthesia (unclear). Self-assessed at home at 0, 1, 2 hours, then hourly for 12 hours or until remedication (diary). Diary returned 5 days later for review and debriefing.	PI (4-point scale) PR (5-point scale) Global rating (5-point scale) Time to remedication	Paracetamol, 600 mg, + codeine, 60 mg (n = 17); paracetamol, 600 mg (n = 22).	At 4 hours paracetamol and combination significantly superior to placebo. Combination also significantly superior to paracetamol for TOTPAR ($p < 0.05$).	Remedication allowed after 2 hours but asked to complete next evaluation before doing so (PI last or baseline score, PR = 0 for all further time points). % remedicating by hour 12: paracetamol 95%; placebo 91%.	88 analysed. Exclusions: 9 did not take medication, 2 remedicated before 2 hours, 1 remedicated with slight pain, 4 did not complete evaluation, 1 took only part of medication, 2 remedicated despite of having some relief from study medication.	None serious reported. Numbers reporting adverse effects (number of effects): paracetamol + codeine 1/17 (1); paracetamol 3/26 (3). NB: includes those reported post-remedication.	5
Forbes, *et al.*, 1990	Impacted third molar (1 or more) n = 269 Age: 15+ years	RCT, double-blind, single then multiple oral dose, 6 parallel groups. General/local anaesthesia (unclear). Self-assessed at home at 0, 1 hours then hourly for 6 hours. Returned 5 days later for review and debriefing.	PI (4-point scale) PR (5-point scale) Pain half gone? (y/n) Global rating (5-point scale) Time to remedication	Paracetamol, 600 mg, + codeine, 60 mg (n = 38; paracetamol, 600 mg (n = 36).	All active medications significantly superior for all measures of total and peak analgesia.	Remedication allowed after 2 hours but asked to complete next evaluation (PI last or baseline score, PR = 0 for all remaining time points). % remedicated by hour 6: placebo 33%; paracetamol 29%.	206 analysed. Exclusions; 3 lost to follow-up, 1 lost report card, 22 did not require medication, 8 remedicated despite having relief from study medication, 6 remedicated with only slight pain, 13 remedicated at < 2 hours, 7 failed to follow instructions, 3 did not complete forms.	None serious reported. Numbers reporting adverse effects (number of effects): paracetamol + codeine 8/40 (9); paracetamol 5/41 (5).	5
Gertzbein, *et al.*, 1986	Postoperative (elective surgery – orthopaedic or general) n = 116 Age range: 16–65 years	RCT, single oral dose, 3 parallel groups. No information on anaesthesia, First oral analgesic given. Interviewed in hospital by nurse-observer at 0, 0.5, 1 hours, then hourly for 5 hours.	PI (5-point scale) PR (5-point scale) 50% PR? (y/n) Global rating (4-point scale): patient and observer VAS PI ('No pain' to 'Worst pain I can imagine')	Paracetamol, 1000 mg, + codeine, 60 mg (n = 45); paracetamol, 1000 mg (n = 45).	Combination seen to give higher efficacy results than paracetamol alone but was not significantly superior for any measure.	Remedication allowed after 1 hour (PI = last score and PR = 0 for all remaining time points). Mean time to remedication: combination 230 minutes; paracetamol 214 minutes.	113 analysed. Exclusions: 1 refused to comply with instructions, 1 vomited within 1 hour of taking study medication, 1 took concomitant analgesia.	None serious reported. Numbers reporting adverse effects (number of effects): paracetamol + codeine 13/47 (13); paracetamol 13/46 (15).	4

continued

TABLE 27 contd *Studies of paracetamol plus codeine versus paracetamol*

Study	Condition and number of patients	Design, study duration and follow-up	Outcome measures	Dosing regimen	Analgesic outcome results	Remedication	Withdrawals and exclusions	Adverse effects	Quality score
Honig & Murray, 1984	Postoperative (elective surgery – abdominal, orthopaedic, rectal, thoracic and vascular) n = 116 Age range: 19–87 years	RCT, single oral dose, 4 parallel groups. No information on anaesthesia, 4-hour washout prior to start. Interviewed in hospital by nurse observer at 0, 0.5, 1 hours then hourly for 6 hours.	PI (4-point scale) PR (5-point scale) 50% PR? (y/n) Global rating (5-point scale)	Paracetamol, 600 mg, + codeine, 60 mg (n = 30); paracetamol, 600 mg (n = 28).	For TOTPAR and global rating both paracetamol and combination significantly superior to placebo ($p < 0.05$). Combination also significantly superior for SPID and number of patients remedicating before 6 hours.	If remedicated, last score used for all further time points. Number not remedicated at 6 hours: placebo 14/30; paracetamol 16/28.	No details given.	None severe except 1 severe dry mouth. Reported in all groups; primarily central nervous system and gastrointestinal effects.	3
Sunshine, *et al.*, 1986	Impacted third molar n = 182 Age: 16+ years	RCT, double-blind, single oral dose, 6 parallel groups. Local anaesthetic (lidocaine/epinephrine). 4-hour washout prior to start. Evaluations at 0, 0.5, 1, 2 and 3 hours in clinic by single observer. At 4, 5 and 6 hours self-assessed. 1 week later met with observer and reviewed forms.	PI (4-point scale) PR (5-point scale) Global rating (4-point scale) Overall improvement (7-point scale) Time to remedication	Paracetamol, 650 mg, + codeine, 60 mg (n = 31); paracetamol, 650 mg (n = 30).	Both paracetamol and combination significantly superior ($p < 0.05$) to placebo for: PID at 1–3 hours, SPID at 4 hours, PR at 2 hours and time to peak effect. Combination also significantly superior for TOTPAR.	Excluded if at < 1 hour. If at > 1 hour, last PI or baseline and PR = 0 used. Full dichotomous data given in Table III. Number remedicating: placebo 13/30; paracetamol 14/30.	182 analysed; no exclusions.	None serious reported; NSD between groups. Numbers reporting adverse effects: paracetamol + codeine 3/31; paracetamol 1/30.	5

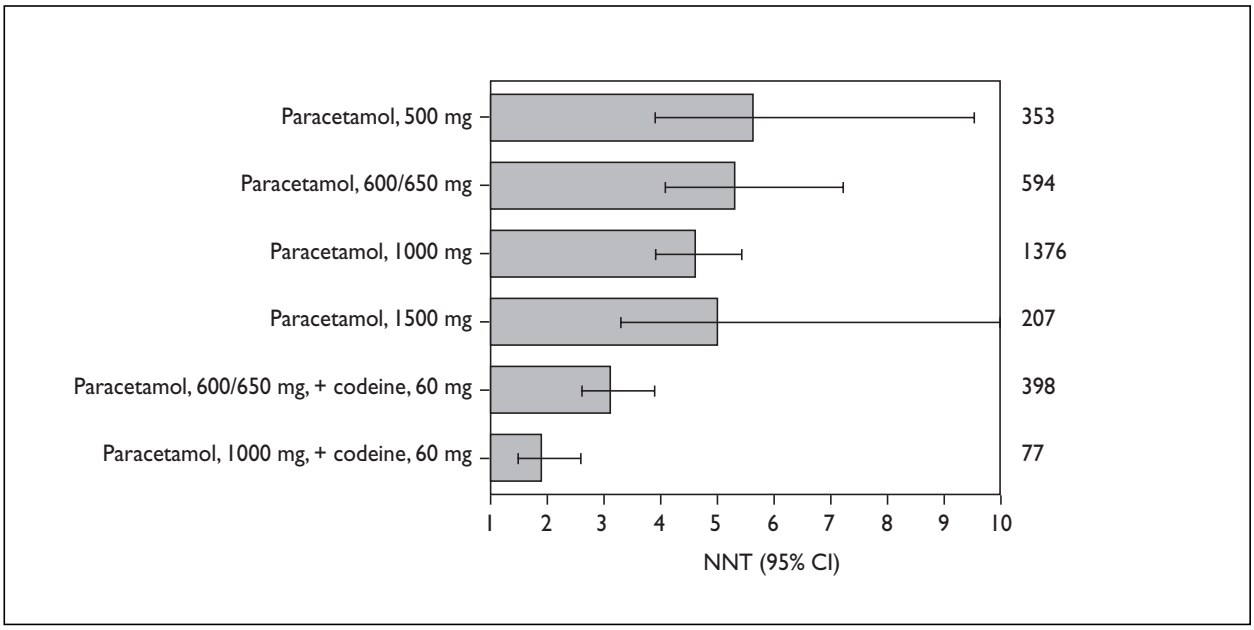


FIGURE 27 *Relative effectiveness of paracetamol doses and paracetamol/codeine combinations*

References

1. Moore A, Collins S, Carroll D, McQuay H. Paracetamol with and without codeine in acute pain: a quantitative systematic review. *Pain* 1997;**70**:193–201.
2. Government Statistical Service. Prescription cost analysis for England 1995. London: Department of Health, 1996.
3. Jadad AR. Meta-analysis of randomised clinical trials in pain relief [DPhil thesis]. Oxford: University of Oxford, 1994.
4. Jadad AR, Carroll D, Moore A, McQuay H. Developing a database of published reports of randomised clinical trials in pain research. *Pain* 1996;**66**:239–46.
5. de Craen AJM, di Giulio G, LampeSchoenmaeckers AJE, Kessels AGH, Kleijnen J. Analgesic efficacy and safety of paracetamol–codeine combinations versus paracetamol alone: a systematic review. *BMJ* 1996;**313**:321–5.
6. Collins SL, Moore RA, McQuay HJ. The visual analogue pain intensity scale: what is moderate pain in millimetres? *Pain* 1997;**72**:95–7.

7. L'Abbé KA, Detsky AS, O'Rourke K. Meta-analysis in clinical research. *Ann Intern Med* 1987;**107**:224–33.

8. Cooper SA. Single-dose analgesic studies: the upside and downside of assay sensitivity. *Adv Pain Res Ther* 1991;**18**:117–24.

9. Moore A, McQuay H, Gavaghan D. Deriving dichotomous outcome measures from continuous data in randomised controlled trials of analgesics. *Pain* 1996;**66**:229–37.

10. Moore A, McQuay H, Gavaghan D. Deriving dichotomous outcome measures from continuous data in randomised controlled trials of analgesics: verification from independent data. *Pain* 1997;**69**:127–30.

11. Moore A, Moore O, McQuay H, Gavaghan D. Deriving dichotomous outcome measures from continuous data in randomised controlled trials of analgesics: use of pain intensity and visual analogue scales. *Pain* 1997;**69**:311–15.

12. DerSimonian R, Laird N. Meta-analysis of clinical trials. *Control Clin Trials* 1986;**7**:177–88.

13. Cook RJ, Sackett DL. The number needed to treat: a clinically useful measure of treatment effect. *BMJ* 1995;**310**:452–4.

14. Berry FN, Miller JM, Levin HM, Bare WW, Hopkinson JH, Feldman AJ. Relief of severe pain with acetaminophen in a new dose formulation versus propoxyphene hydrochloride 65 mg and placebo: a comparative double blind study. *Curr Ther Res* 1975;**17**:361–8.

15. Dolci G, Ripari M, Pacifici L, Umile A. Evaluation of piroxicam-beta-cyclodextrin, piroxicam, paracetamol and placebo in post-operative oral surgery pain. *Int J Clin Pharmacol Res* 1994;**14**:185–91.

16. Kiersch TA, Halladay SC, Hormel PC. A single-dose, double-blind comparison of naproxen sodium, acetaminophen, and placebo in postoperative dental pain. *Clin Ther* 1994;**16**:394–404.

17. Turek M, Baird W. Double blind parallel comparison of ketoprofen (Orudis), acetaminophen plus codeine, and placebo in postoperative pain. *J Clin Pharmacol* 1988;**28**:S23–8.

18. McQuay HJ, Tramèr M, Nye BA, Carroll D, Wiffen PJ, Moore RA. A systematic review of antidepressants in neuropathic pain. *Pain* 1996;**68**:217–27.

19. Tramèr M, Moore A, McQuay H. Prevention of vomiting after paediatric strabismus surgery: a systematic review using the numbers-needed-to-treat method. *Br J Anaesth* 1995;**75**:556–61.

20. Soll JC, McQueen MC. Respiratory distress syndrome. In: Sinclair JC, Bracken ME, editors. Effective care of the newborn infant. Oxford: Oxford University Press, 1992: chapter 15, p. 333.

21. Moore RA, McQuay HJ. Single-patient data meta-analysis of 3453 postoperative patients: oral tramadol versus placebo, codeine and combination analgesics. *Pain* 1997;**69**:287–94.

Studies included

Beaver WT, McMillan D, 1980. Methodological considerations in the evaluation of analgesic combinations: acetaminophen (paracetamol) and hydrocodone in postpartum pain. *Br J Clin Pharmacol*;**10** suppl:215–23S.

Bentley KC, Head TW, 1987. The additive analgesic efficacy of acetaminophen, 1000 mg, and codeine, 60 mg, in dental pain. *Clin Pharmacol Ther*;**42**:634–40.

Berry FN, Miller JM, Levin HM, Bare WW, Hopkinson JH, Feldman AJ, 1975. Relief of severe pain with acetaminophen in a new dose formulation versus propoxyphene hydrochloride 65 mg and placebo: a comparative double blind study. *Curr Ther Res*;**17**:361–8.

Bjune K, Stubhaug A, Dodgson MS, Breivik H, 1996. Additive analgesic effect of codeine and paracetamol can be detected in strong, but not moderate, pain after Caesarean section. Baseline pain-intensity is a determinant of assay-sensitivity in a postoperative analgesic trial [see comments]. *Acta Anaesthesiol Scand*;**40**:399–407.

Cooper SA, Kupperman A, 1991. The analgesic efficacy of flurbiprofen compared to acetaminophen with codeine. *J Clin Dent*;**2**:70–4.

Cooper SA, Precheur H, Rauch D, Rosenheck A, Ladov M, Engel J, 1980. Evaluation of oxycodone and acetaminophen in treatment of postoperative dental pain. *Oral Surg Oral Med Oral Pathol*;**50**:496–501.

Cooper SA, Breen JF, Giuliani RL, 1981. The relative efficacy of indoprofen compared with opioid analgesic combinations. *J Oral Surg*;**39**:21–5.

Cooper SA, Erlichman MC, Mardirossian G, 1986. Double blind comparison of an acetaminophen codeine caffeine combination in oral surgery pain. *Anesth Prog*;**33**:139–42.

Cooper SA, Firestein A, Cohn P, 1988. Double blind comparison of meclofenamate sodium with acetaminophen, acetaminophen with codeine and placebo for relief of postsurgical dental pain. *J Clin Dent*;**1**:31–4.

Cooper SA, Schachtel BP, Goldman E, Gelb S, Cohn P, 1989. Ibuprofen and acetaminophen in the relief of acute pain: a randomized, double blind, placebo controlled study. *J Clin Pharmacol*;**29**:1026–30.

Desjardins PJ, Cooper SA, Finizio T, 1986. Efficacy of low dose combination analgesics: acetaminophen/codeine, aspirin/butalbital/caffeine/codeine, and placebo in oral surgery pain. *Anesth Prog*;**33**:143–6.

Dionne RA, Snyder J, Hargreaves KM, 1994. Analgesic efficacy of flurbiprofen in comparison with acetaminophen, acetaminophen plus codeine, and placebo after impacted third molar removal. *J Oral Maxillofac Surg*;**52**:919–24; discussion 25–6.

Dolci G, Ripari M, Pacifici L, Umile A, 1994. Evaluation of piroxicam-beta-cyclodextrin, piroxicam, paracetamol and placebo in post-operative oral surgery pain. *Int J Clin Pharmacol Res*;**14**:185–91.

Fassolt A, Stocker H, 1983. Behandlung des post-operativen Wundschmerzes mit Suprofen [Treatment of postoperative wound pain with suprofen]. *Arzneimittelforschung*;**33**:1327–30.

Forbes JA, Beaver WT, White EH, White RW, Neilson GB, Shackleford RW, 1982. Diflunisal: a new oral analgesic with an unusually long duration of action. *JAMA*;**248**:2139–42.

Forbes JA, Kolodny AL, Beaver WT, Shackleford RW, Scarlett VR, 1983. A 12 hour evaluation of the analgesic efficacy of diflunisal, acetaminophen, and acetaminophen codeine combination, and placebo in postoperative pain. *Pharmacotherapy*;**3**:47–54S.

Forbes JA, Barkaszi BA, Ragland RN, Hankle JJ, 1984. Analgesic effect of acetaminophen, phenyltoloxamine and their combination in postoperative oral surgery pain. *Pharmacotherapy*;**4**:221–6.

Forbes JA, Kolodny AL, Chachich BM, Beaver WT, 1984. Nalbuphine, acetaminophen, and their combination in postoperative pain. *Clin Pharmacol Ther*;**35**:843–51.

Forbes JA, Jones KF, Smith WK, Gongloff CM, 1986. Analgesic effect of an aspirin codeine butalbital caffeine combination and an acetaminophen codeine combination in postoperative oral surgery pain. *Pharmacotherapy*;**6**:240–7.

Forbes JA, Butterworth GA, Burchfield WH, Yorio CC, Selinger LR, Rosenmertz SK, *et al.,* 1989. Evaluation of flurbiprofen, acetaminophen, an acetaminophen-codeine combination, and placebo in postoperative oral surgery pain. *Pharmacotherapy*;**9**:322–30.

Forbes JA, Butterworth GA, Burchfield WH, Beaver WT, 1990. Evaluation of ketorolac, aspirin, and an acetaminophen–codeine combination in postoperative oral surgery pain. *Pharmacotherapy*;**10**:77–93S.

Forbes JA, Kehm CJ, Grodin CD, Beaver WT, 1990. Evaluation of ketorolac, ibuprofen, acetaminophen, and an acetaminophen–codeine combination in postoperative oral surgery pain. *Pharmacotherapy*;**10**:94–105S.

Forbes JA, Bates JA, Edquist IA, Burchfield WH, Smith FG, Schwartz MK, *et al.,* 1994. Evaluation of two opioid-acetaminophen combinations and placebo in post-operative oral surgery pain. *Pharmacotherapy*;**14**:139–46.

Gertzbein SD, Tile M, McMurty RY, Kellam JF, Hunter GA, Keith RG, *et al.,* 1986. Analysis of the analgesic efficacy of acetaminophen 1000 mg, codeine phosphate 60 mg, and the combination of acetaminophen 1000 mg and codeine phosphate 60 mg in the relief of post-operative pain. *Pharmacotherapy*;**6**:104–7.

Heidrich G, Slavic Svircev V, Kaiko RF, 1985. Efficacy and quality of ibuprofen and acetaminophen plus codeine analgesia. *Pain*;**22**:385–97.

Honig S, Murray KA, 1984. An appraisal of codeine as an analgesic: single dose analysis. *J Clin Pharmacol*;**24**:96–102.

Jain AK, Ryan JR, McMahon FG, Smith G, 1986. Comparison of oral nalbuphine, acetaminophen, and their combination in postoperative pain. *Clin Pharmacol Ther*;**39**:295–9.

Kiersch TA, Halladay SC, Hormel PC, 1994. A single-dose, double-blind comparison of naproxen sodium, acetaminophen, and placebo in postoperative dental pain. *Clin Ther*;**16**:394–404.

Laska EM, Sunshine A, Zighelboim I, Roure C, Marrero I, Wanderling J, *et al.,* 1983. Effect of caffeine on acetaminophen analgesia. *Clin Pharmacol Ther*;**33**:498–509.

Lehnert S, Reuther J, Wahl G, Barthel K, 1990. Wirksamkeit von Paracetamol (Tylenol) und Acetylsalizylsaure (Aspirin) bei postoperativen Schmerzen [The efficacy of paracetamol (Tylenol) and acetyl salicylic acid (Aspirin) in treating postoperative pain]. *Dtsch Zahnarztl Z*;**45**:23–6.

McQuay HJ, Carroll D, Frankland T, Harvey M, Moore A, 1990. Bromfenac, acetaminophen, and placebo in orthopedic postoperative pain. *Clin Pharmacol Ther*;**47**:760–6.

Mehlisch DR, Frakes LA, 1984. A controlled comparative evaluation of acetaminophen and aspirin in the treatment of postoperative pain. *Clin Ther*;**7**:89–97.

Mehlisch DR, Sollecito WA, Helfrick JF, Leibold DG, Markowitz R, Schow CEJ, *et al.,* 1990. Multicenter clinical trial of ibuprofen and acetaminophen in the treatment of postoperative dental pain. *J Am Dent Assoc*;**121**:257–63.

Mehlisch DR, Jasper RD, Brown P, Korn SH, McCarroll K, Murakami AA, 1995. Comparative study of ibuprofen lysine and acetaminophen in patients with postoperative dental pain. *Clin Ther*;**17**:852–60.

Pande AC, Pyke RE, Greiner M, Cooper SA, Benjamin R, Pierce MW, 1996. Analgesic efficacy of the kappa-receptor agonist, enadoline, in dental surgery pain. *Clin Neuropharmacol*;**19**:92–7.

Petti A, 1985. Postoperative pain relief with pentazocine and acetaminophen: comparison with other analgesic combinations and placebo. *Clin Ther*;**8**:126–33.

Rubin A, Winter LJ, 1984. A double blind randomized study of an aspirin/caffeine combination versus acetaminophen/aspirin combination versus acetaminophen versus placebo in patients with moderate to severe post partum pain. *J Int Med Res*;**12**:338–45.

Schachtel B, Thoden W, Baybutt R, 1989. Ibuprofen and acetaminophen in the relief of postpartum episiotomy pain. *J Clin Pharmacol*;**29**:550–3.

Seymour RA, Kelly PJ, Hawkesford JE, 1996. The efficacy of ketoprofen and paracetamol (acetaminophen) in postoperative pain after third molar surgery. *Br J Clin Pharmacol*;**41**:581–5.

Stubhaug A, Grimstad J, Breivik H, 1995. Lack of analgesic effect of 50 and 100 mg oral tramadol after orthopaedic surgery: a randomized, double-blind, placebo and standard active drug comparison. *Pain*;**62**:111–18.

Sunshine A, Marrero I, Olson N, McCormick N, Laska EM, 1986. Comparative study of flurbiprofen, zomepirac sodium, acetaminophen plus codeine, and acetaminophen for the relief of postsurgical dental pain. *Am J Med*;**80**:50–4.

Sunshine A, Zighelboim I, De Castro A, Sorrentino JV, Smith DS, Bartizek RD, *et al.,* 1989. Augmentation of acetaminophen analgesia by the antihistamine phenyltoloxamine. *J Clin Pharmacol*;**29**:660–4.

Sunshine A, Olson NZ, Zighelboim I, de Castro A, 1993. Ketoprofen, acetaminophen plus oxycodone, and acetaminophen in the relief of postoperative pain. *Clin Pharmacol Ther*;**54**:546–55.

Turek M, Baird W, 1988. Double blind parallel comparison of ketoprofen (Orudis), acetaminophen plus codeine, and placebo in postoperative pain. *J Clin Pharmacol*;**28**:S23–8.

Winnem B, Samstad B, Breivik H, 1981. Paracetamol, tiaramide and placebo for pain relief after orthopedic surgery. *Acta Anaesthesiol Scand*;**25**:209–14.

Winter LJ, Appleby F, Ciccone PE, Pigeon JG, 1983. A comparative study of an acetaminophen analgesic combination and aspirin in the treatment of postoperative oral surgery pain. *Curr Ther Res*;**33**:115–22.

Young RE, Quigley JJ, Archambault WAJ, Gordon LL. Butorphanol/acetaminophen double blind study in postoperative pain. *J Med* 1979;**10**:239–56.

Chapter 9

Oral ibuprofen and diclofenac in postoperative pain

Summary

The aim of this review was to compare ibuprofen and diclofenac in postoperative pain. Studies were identified by an extensive literature search, with additional reports being identified from the reference lists of reports, review articles and textbooks.

The studies of interest were randomised, controlled, single dose comparisons of ibuprofen or diclofenac against placebo. Summed pain relief or SPID over 4–6 hours was extracted, and converted into dichotomous information yielding the numbers of patients with at least 50% pain relief. This was then used to calculate the relative benefit and the NNT for one patient to achieve at least 50% pain relief.

In all, 34 reports compared ibuprofen and placebo (3591 patients), six compared diclofenac with placebo (840 patients), and there were two direct comparisons of diclofenac, 50 mg, and ibuprofen, 400 mg, (130 patients). In postoperative pain, ibuprofen, 200 mg, had an NNT of 3.3 (95% CI, 2.8–4.0) compared with placebo, ibuprofen, 400 mg, had an NNT of 2.7 (95% CI, 2.5–3.0) and ibuprofen, 600 mg, had an NNT of 2.4 (95% CI, 1.9–3.3). Diclofenac, 50 mg, had an NNT of 2.3 (95% CI, 2.0–2.7) compared with placebo in established postoperative pain; diclofenac, 100 mg, had an NNT of 1.8 (95% CI, 1.5–2.1).

When diclofenac, 50 mg, was compared directly with ibuprofen, 400 mg, there was no significant difference between the two treatments. Ibuprofen showed a clear dose response with relative efficacy similar to that for diclofenac. Both drugs work well. Choosing between them is an issue of dose, safety and cost.

This chapter of the review has been published in full by Collins and colleagues.[1]

Introduction

Ibuprofen and diclofenac are two of the most widely used NSAIDs, with ibuprofen commonly available without prescription. In England in 1996, ibuprofen accounted for nearly 5.5 million prescriptions (31% of total NSAID prescriptions) and diclofenac for nearly 6 million prescriptions (34%), although it is not known how many were for acute pain conditions.[2]

With an increasing amount of surgery being performed as day cases, it is important to know which drug should be recommended for postoperative pain relief. The relative efficacy of the two drugs was compared to allow a balanced decision to be made, based on efficacy, safety and cost.

Methods

Single dose, RCTs of ibuprofen and diclofenac in postoperative pain (post-dental extraction, post-surgical or postpartum pain) were sought. Different search strategies were used to identify eligible reports from MEDLINE (1966–December 1996), EMBASE (1980–January 1997), the Cochrane Library (August 1996), Biological Abstracts (1985–96) and the Oxford Pain Relief Database (1950–94).[3] A search was undertaken for each drug using the terms, 'clinical trial', 'trial', 'study', 'random*', 'double blind', 'analgesi*' and 'pain*', together with 'ibuprofen', 'Brufen', 'propionic acid' and 'isobutylphenyl propionic acid' for the ibuprofen search and, for the diclofenac search, 'diclofenac', and 76 brand names.[4] Each was a broad free text search, including various combinations of the words, and without restriction to language. Additional reports were identified from the reference lists of retrieved reports, review articles and textbooks.

Included reports

Neither pharmaceutical companies nor authors of papers were contacted for unpublished reports. Abstracts and review articles were not considered. The inclusion criteria used were:

- randomised allocation to treatment groups which included ibuprofen or diclofenac and placebo
- full journal publication
- established postoperative pain with the pain

outcome measured using a five-point pain relief scale with standard wording (none, slight, moderate, good, complete) or a four-point pain intensity scale (none, mild, moderate or severe) or a VAS for pain relief or pain intensity, TOTPAR or SPID (at 4, 5 or 6 hours) as a derived pain relief outcome (or sufficient data provided to allow their calculation)
- postoperative oral administration for ibuprofen and postoperative oral, rectal, intravenous or intramuscular administration for diclofenac
- adult patients
- baseline pain of moderate to severe intensity (for VAS this equates to > 30 mm[5])
- double-blind design.

Excluded reports

Reports were excluded of:

- ibuprofen or diclofenac used for the relief of other pain conditions
- controlled release formulations
- ibuprofen or diclofenac used in combination with other drugs
- trials which reported data from a crossover design as a single data set
- trials in which the number of patients per treatment group was less than ten[6]
- trials which included pain relief data collected after additional analgesia was given.

Data extraction and analysis

From each report the numbers of patients treated was taken, together with the mean TOTPAR, SPID, VAS TOTPAR or VAS SPID, study duration and the doses given. Information on adverse events was also extracted. For each report, the mean TOTPAR, SPID, VAS TOTPAR or VAS SPID values for active and placebo were converted to % maxTOTPAR or % maxSPID by division into the calculated maximum value.[7] The proportion of patients in each treatment group who achieved at least 50% maxTOTPAR was calculated using verified equations.[8–10] The proportions were then converted into the number of patients achieving at least 50% maxTOTPAR by multiplying by the total number of patients in the treatment group. Information on the number of patients with at least 50% maxTOTPAR for active drug and placebo was then used to calculate the relative benefit and NNT.

Relative benefit estimates were calculated with 95% CIs using a random effects model;[11] the random effects model was chosen because it produces the most conservative estimate (homogeneity was assumed when $p > 0.1$). NNT and 95% CIs were calculated using Cook and Sackett's method.[12] A statistically significant difference from control was assumed when the 95% CI of the relative benefit did not include 1.

The number of patients experiencing at least 50% pain relief with placebo (the CER) can vary greatly with the relatively small sample sizes used typically in analgesic trials affecting the apparent efficacy of an analgesic.[13] To allow for this variation, the relative benefit and NNT for each dose of ibuprofen and diclofenac were also calculated using a fixed CER of 19%. This value was obtained from data for 4378 patients given placebo, pooled from 124 single dose analgesic trials meeting identical inclusion criteria included in this and similar systematic reviews (843/4378 patients experienced at least 50% pain relief).[14–17]

Results

Ibuprofen versus placebo

There were 34 reports of 35 trials that fulfilled our inclusion criteria; a total of 2214 patients were given ibuprofen and 1377 placebo. Details of the studies are presented in *Table 28* with their references being listed at the end of this chapter. The author of one report was contacted for information on the number of patients in each treatment arm.[18]

Oral surgery pain was investigated in 25 trials (predominantly third molar extraction with bone removal), in five trials postpartum pain was investigated (predominantly episiotomy and Caesarean section), and in a further four trials postoperative pain was studied (one tonsillectomy, one inguinal hernia, one orthopaedic surgery and one general surgery). The doses of ibuprofen prescribed were 50 mg in one trial, 100 mg in two, 200 mg in eight, 400 mg in 30, 600 mg in three, and 800 mg in one.

The CER (the proportion of patients given placebo experiencing at least 50% pain relief) ranged from 0% to 67% (median 12%) *(Figure 28)*. For the single trial of ibuprofen, 50 mg, the EER (the proportion of patients given ibuprofen experiencing at least 50% pain relief) was 28%. For ibuprofen, 100 mg, two trials gave EER values of 27% and 8%. For ibuprofen, 200 mg, the EER varied between 6% and 57% (median 39%), and for ibuprofen, 400 mg, it varied between 13% and 100% (median 60%). For the single trial of ibuprofen, 800 mg, the EER was 100%. The 100 mg and 200 mg data sets were homogeneous but the 400 mg and 600 mg data sets were not.

TABLE 28 *Studies of ibuprofen versus placebo*

Study	Condition and number of patients	Design, study duration and follow-up	Outcome measures	Dosing regimen	Analgesic outcome results	Remedication	Withdrawals and exclusions	Adverse effects
Ahlstrom, *et al.*, 1993	Third molar extraction n = 127 Age range: 18–40 years	RCT, double-blind, single oral dose, parallel groups. 4-hour washout before start. Evaluated at 0, 20, 40 minutes 1 hour then hourly intervals for 6 hours. Medication taken when baseline PI at least moderate (> 30 mm).	VAS PI: 'no pain at all' to 'agonising pain' Global rating by patient	Ibuprofen, 400 mg (n = 32); placebo (n = 30).	Ibuprofen significantly superior to placebo by 40 minutes (p = 0.01); this continued for 6 hours. TOTPI & SPID: ibuprofen significantly superior to placebo (p < 0.0001). 6 hour SPID: ibuprofen 188 mm; placebo 32 mm.	Patients allowed to remedicate after 1 hour. After remedication, PI = last score carried forward for all further time points.	97 analysed. Exclusions: 30 for various protocol violations.	None serious reported and no patient withdrew as a result. Numbers reporting adverse effects (number of effects): ibuprofen 3/32 (?); placebo 2/30 (?).
Arnold, *et al.*, 1990	General surgery (including gynaecological and orthopaedic) n = 59 Age range: 22–70 years	RCT, double-blind, single oral dose, parallel groups. Assessed by single nurse-observer at 0, 0.5, 1 hour then hourly intervals for 6 hours. Medication taken when baseline PI moderate to severe.	PI (standard 4-point scale) PR (5-point scale), viz, standard time to meaningful relief Global rating (5-point scale) by patient	Ibuprofen, 400 mg (n = 15); placebo (n = 14).	Ibuprofen not significantly superior to placebo for either SPID or TOTPAR. 6-hour TOTPAR: ibuprofen 4.2; placebo 1.5.	After remedication PR = 0 and PI = baseline score for all further time points.	No information on any exclusions.	Difference in occurrence not significant between groups. No patient withdrew from either ibuprofen or placebo group as a result of adverse events.
Bakshi, *et al.*, 1994	Third molar extraction n = 257 Age: adults up to 65 years	RCT, double-blind, single oral dose, parallel groups. Local anaesthetic. Self-assessed at 0, 20 and 40 minutes, 1, 1.5, and 2 hours, then hourly for 6 hours. Medication taken when baseline PI was at least severe.	VAS PI: 'no pain' to 'pain could not be worse' PR (5-point scale): none, poor, moderate, sufficient, total Global rating (5-point scale) by patient and by observer	Ibuprofen, 400 mg (n = 80); placebo (n = 82).	Ibuprofen significantly superior to placebo for TOTPAR and both global ratings (p < 0.01). 6-hour TOTPAR: ibuprofen 14.9; placebo 8.85.	Patients allowed to remedicate after 1 hour. If they remedicated earlier, data excluded from efficacy analysis. After remedication PR = 0 and PI = last score for all further time points.	245 analysed. Exclusions: 9 did not experience severe pain, 2 remedicated before 1 hour, 1 completed diary incorrectly.	None serious reported and no patient withdrew as a result. Numbers reporting adverse effects (number of effects): ibuprofen 6/80 (?); placebo 5/82 (?)
Cooper, *et al.*, 1977	Third molar extraction n = 245 Age: ?	RCT, double-blind, single oral dose, parallel groups. Local anaesthetic. Self-assessed at home at 0 hours then hourly for 4 hours. Medication taken when baseline PI moderate to severe.	PI (standard 4-point scale) PR (standard 5-point scale) 50% PR? (y/n) Global rating (5-point scale) by patient	Ibuprofen, 400 mg (n = 40); ibuprofen, 200 mg (n = 38); placebo (n = 40).	Ibuprofen at both doses significantly superior to placebo for all measures of efficacy (p < 0.05). 4-hour TOTPAR: ibuprofen, 400 mg, 7.32; ibuprofen, 200 mg, 6.27; placebo 3.32.	Patients allowed to remedicate after 2 hours. If they remedicated earlier data excluded from efficacy analysis. After remedication PR = 0 and PI = baseline score for all further time points.	192 analysed. Exclusions: 17 provided uninterpretable data, 12 took confounding medication, 10 lost to follow-up, 9 did not need medication, 5 fell asleep.	None serious reported and no patient withdrew as a result. No individual data provided but NSD in occurrence between groups.
Cooper, *et al.*, 1982	Third molar extraction n = 316 Age range: 16–65 years	RCT, double-blind, single oral dose, parallel groups. Mostly local anaesthetic. Self-assessed at home at 0 hours then hourly for 4 hours. Medication taken when baseline PI was moderate to severe.	PI (standard 4-point scale) PR (standard 5-point scale) 50% PR? (y/n) Global rating (5-point scale) by patient	Ibuprofen, 400 mg (n = 38); placebo (n = 46); ibuprofen, 400 mg, + codeine, 60 mg (n = 41); codeine, 60 mg (n = 41).	All active treatments significantly superior to placebo for SPID and TOTPAR (no p-value given). 4-hour TOTPAR: ibuprofen 8.39; placebo 2.65; ibuprofen + codeine 9.39; codeine 4.12.	Patients allowed to remedicate after 1 hour. If they remedicated earlier data excluded from efficacy analysis. After remedication PR = 0 and PI = baseline score for all further time points.	249 analysed. Exclusions: 30 lost to follow-up, 15 did not require medication, 11 remedicated before 1 hour, 6 missed more than 1 evaluation, 3 medicated with slight pain, 1 did not take all medication, 1 medicated over 24 hours after surgery.	None serious reported; no patient withdrew as a result. Numbers reporting adverse effects (number of effects): ibuprofen 11/38 (12); placebo 5/46 (6); ibuprofen + codeine 18/41 (20); codeine 11/41 (11).
Cooper, *et al.*, 1988	Third molar extraction n = 201 Age: Adult	RCT, double-blind, single oral dose, parallel groups. Local anaesthetic + sedative. Self-assessed at home at 0 hours, then hourly for 6 hours. Medication taken when baseline PI was moderate to severe.	PI (standard 4-point scale) PR (standard 5-point scale) 50% PR? (y/n) Global rating (5-point scale) by patient	Ibuprofen, 400 mg (n = 37); placebo (n = 43).	Ibuprofen significantly superior to placebo for all measures of efficacy (p < 0.01). 6-hour TOTPAR: ibuprofen 11.32; placebo 4.67.	Patients allowed to remedicate after 1 hour. If they remedicated earlier data excluded from efficacy analysis. After remedication PR = 0 and PI = baseline score for all further time points.	161 analysed. Exclusions: 20 did not require medication, 13 lost to follow-up, 7 for various protocol violations.	None serious reported; no patient withdrew as a result. Numbers reporting adverse effects (number of effects): ibuprofen 10/40 (14); placebo 7/45 (7).

continued

TABLE 28 contd *Studies of ibuprofen versus placebo*

Study	Condition and number of patients	Design, study duration and follow-up	Outcome measures	Dosing regimen	Analgesic outcome results	Remedication	Withdrawals and exclusions	Adverse effects
Cooper, *et al.*, 1989	Third molar extraction n = 194 Age: 16+ years	RCT, double-blind, single oral dose, parallel groups. Local anaesthetic. Self-assessed at home at 0, 0.5, 1 hours then hourly for 6 hours. Medication taken when baseline PI moderate to severe.	PI (standard 4-point scale) PR (standard 5-point scale) 50% PR? (y/n) Global rating (5-point scale) by patient	Ibuprofen, 400 mg (n = 61); placebo (n = 64).	Ibuprofen significantly superior to placebo for all measures of efficacy ($p < 0.0001$). 6-hour TOTPAR: ibuprofen 11.32; placebo 4.67.	Patients allowed to remedicate after 1 hour. If they remedicated earlier data excluded from efficacy analysis. No details on how data were handled for remedication.	184 analysed. Exclusions: 2 lost to follow-up, 2 did not require medication, 4 missed more than 1 evaluation, 1 had insufficient baseline pain, 1 failed to complete diary at appropriate time.	None serious reported; no patient withdrew as a result. Numbers reporting adverse effects (number of effects): ibuprofen 5/63 (6); placebo 7/64 (7).
Forbes, *et al.*, 1984	Third molar extraction n = 136 Age: 15+ years	RCT, double-blind, single oral dose, parallel groups. General anaesthetic. Self-assessed at home at 0 hours, then hourly for 12 hours. Medication taken when baseline PI moderate to severe. Follow-up 5 days post-surgery with research nurse.	PI (standard 4-point scale) PR (standard 5-point scale) 50% PR? (y/n) Global rating (5-point scale) by patient	Ibuprofen, 400 mg (n = 28); placebo (n = 28).	Ibuprofen significantly superior to placebo for all measures of efficacy ($p < 0.01$). 6-hour TOTPAR: ibuprofen 15.79; placebo 3.79. 4-hour TOTPAR: ibuprofen 10.75; placebo 2.79.	Patients allowed to remedicate after 2 hours. If they remedicated earlier data excluded from efficacy analysis. After remedication PR = 0 and PI = baseline or last score (whichever was greater) for all further time points.	109 analysed. Exclusions: 21 did not require medication, 2 took rescue medication instead of trial medication, 2 remedicated despite having some relief, 2 remedicated before 2 hours.	None serious reported; no patient withdrew as a result. Numbers reporting adverse effects (number of effects): ibuprofen 5/28 (6); placebo 3/28 (3).
Forbes, *et al.*, 1990	Third molar extraction n = 269 Age: 15+ years	RCT, double-blind, single then multiple oral dose, parallel groups. General/ local anaesthetic (unclear). Self-assessed at home at 0, 1 hours then hourly for 6 hours. Medication taken when baseline PI moderate to severe. Follow-up 5 days postsurgery.	PI (standard 4-point scale) PR (standard 5-point scale) 50% PR? (y/n) Global rating (5-point scale) by patient	Ibuprofen, 400 mg (n = 32); placebo (n = 34).	Ibuprofen significantly superior to placebo for all measures of analgesia ($p < 0.05$ at least). 6-hour TOTPAR: ibuprofen 10.47; placebo 1.88.	Patients allowed to remedicate after 2 hours. If they remedicated earlier data excluded from efficacy analysis. After remedication PR = 0 and PI = baseline or last score (whichever was greater) for all further time points.	206 analysed. Exclusions: 3 lost to follow-up, 1 lost report card, 22 did not require medication, 8 remedicated despite having relief from study medication, 6 remedicated with only slight pain, 13 remedicated at < 2 hours, 7 failed to follow instructions, 3 did not complete forms.	None serious reported; no patient withdrew as a result. Numbers reporting adverse effects (number of effects): ibuprofen 8/43 (9); placebo 0/38 (0).
Forbes, *et al.*, 1991	Third molar extraction n = 395 Age: 15+ years	RCT, double-blind, single oral dose, parallel groups. Local anaesthetic. 4 hour caffeine washout prior to start. Self-assessed at 0, 0.5, 1 hours then hourly for 8 hours. Medication taken when baseline PI moderate to severe. Follow-up 5 days postsurgery. Multicentre (2 sites).	PI (standard 4-point scale) PR (standard 5-point scale) 50% PR? (y/n) Global rating (5-point scale) by patient	Ibuprofen, 50 mg (n = 57); ibuprofen, 100 mg (n = 49); ibuprofen, 200 mg (n = 48); placebo (n = 51); ibuprofen, 100 mg, + caffeine, 100 mg (n = 49); ibuprofen, 200 mg, + caffeine, 100 mg (n = 44).	All ibuprofen treatments significantly superior to placebo for all measures ($p < 0.05$ at least). 8-hour TOTPAR: ibuprofen, 50 mg, 8.82; ibuprofen, 100 mg, 8.46; ibuprofen, 200 mg, 10.00; placebo 2.58; 100 mg combination 11.29; 200 mg combination 15.58 (6-hour TOTPAR calculated from mean hourly scores).	Patients allowed to remedicate after 2 hours. If they remedicated earlier data excluded from efficacy analysis. After remedication PR = 0 and PI = baseline or last score (whichever was greater) for all further time points.	298 analysed. Exclusions: 33 did not require medication, 14 remedicated at < 2 hours, 1 ate caffeine-containing food, 2 medicated for headache, 1 rated only one side of mouth, 1 form completed by relative, 3 lacked consistency, 22 evaluated at incorrect time, 3 incomplete forms.	None serious reported; no patient withdrew as a result. Numbers reporting adverse effects (number of effects): ibuprofen, 50 mg, 10/63 (15); ibuprofen, 100 mg, 5/62 (6); ibuprofen, 200 mg, 6/60 (6); placebo 8/61 (8); 100 mg caffeine combination 12/58 (15); 200 mg caffeine combination 8/58 (9).
Forbes, *et al.*, 1991	Third molar extraction n = 288 Age: 15+ years	RCT, double-blind, single oral dose, parallel groups. Local anaesthetic. Self-assessed at home at 0, 1 hours then hourly for 8 hours. Medication taken when baseline PI moderate to severe. Follow-up 5 days postsurgery.	PI (standard 4-point scale) PR (standard 5-point scale) 50% PR? (y/n) Global rating (5-point scale) by patient	Ibuprofen, 400 mg (n = 37); placebo (n = 39).	Ibuprofen significantly superior to placebo for all measures of efficacy ($p < 0.05$ at least). 8-hour TOTPAR: ibuprofen 14.30; placebo 2.59. 6-hour TOTPAR (calculated from mean hourly scores): ibuprofen 10.97; placebo 2.49.	Patients allowed to remedicate after 2 hours. If they remedicated earlier data excluded from efficacy analysis. After remedication PR = 0 and PI = baseline or last score (whichever was greater) for all further time points.	241 analysed. Exclusions: 7 lost to follow-up, 12 did not require medication, 4 remedicated with some relief, 1 re-medicated with slight pain, 19 remedicated before 2 hours, 2 lacked consistency, 1 did not complete form, 1 took only part of medication.	None serious reported; no patient withdrew as a result. Numbers reporting adverse effects (number of effects): ibuprofen 7/43 (8); placebo 3/47 (3).

continued

TABLE 28 contd *Studies of ibuprofen versus placebo*

Study	Condition and number of patients	Design, study duration and follow-up	Outcome measures	Dosing regimen	Analgesic outcome results	Remedication	Withdrawals and exclusions	Adverse effects
Forbes, *et al.*, 1992	Third molar extraction n = 338 Age: 15+ years	RCT, double-blind, single oral dose, parallel groups. Local anaesthetic. Self-assessed at home at 0 and 1 hour, then hourly for 8 hours. Medication taken when baseline PI moderate to severe. Follow-up 5 days postsurgery.	PI (standard 4-point scale) PR (standard 5-point scale) 50% PR? (y/n) Global rating (5-point scale) by patient	Ibuprofen, 400 mg (n = 38); placebo (n = 38).	Ibuprofen significantly superior to placebo for all measures of efficacy ($p < 0.01$). 8-hour TOTPAR: ibuprofen 14.82; placebo 2.34. 6-hour TOTPAR (calculated from mean hourly scores): ibuprofen 11.79; placebo 2.06.	Patients allowed to remedicate after 2 hours. If they remedicated earlier data was excluded from efficacy analysis. After remedication PR = 0 and PI = baseline or last score (whichever was greater) for all further time points.	280 analysed. Exclusions; 3 did not return form, 14 did not require medication, 4 re-medicated despite some relief, 6 remedicated with slight pain, 18 remedicated before 2 hours, 2 lacked consistency, 2 did not complete form, 2 took only part of medication, 5 took back-up medication.	None serious reported; no patient withdrew as a result. Numbers reporting adverse effects (number of effects): ibuprofen 4/45 (8); placebo 2/46 (5).
Frame, *et al.*, 1989	Third molar extraction n = 148 Age: 16+ years	RCT, double-blind, single oral dose, parallel groups. Local anaesthetic. Self-assessed at home at 0, 0.5, and 1 hours, then hourly for 5 hours. Medication taken when baseline PI at least moderate.	PI (non-standard 9-point scale) PR (standard 5-point scale) 50% PR? (y/n)	Ibuprofen, 400 mg (n = 42); placebo (n = 38).	At 2 and 3 hours ibuprofen significantly superior to placebo. 5-hour TOTPAR (calculated from graph): ibuprofen 12.85; placebo 7.95.	Patients allowed to remedicate after 2 hours. No information provided on how data were handled for patients who remedicated.	123 analysed. Exclusions: 9 did not take medication, 7 lost to follow-up, 1 was asleep so did not complete form, 1 had complications so did not complete form, 7 had slight pain.	None serious reported; no patient withdrew as a result. Numbers reporting adverse effects (number of effects): ibuprofen 2/42 (2); placebo 1/38 (3).
Fricke, *et al.*, 1993	Third molar extraction n = 207 Age: 15+ years	RCT, double-blind, single oral dose, parallel groups. 72-hour washout prior to start. Local anaesthetic. Self-assessed at home at 0, 20, 30, 40, and 60 minutes, then hourly for 12 hours. Medication taken when baseline PI moderate. Review 1/2 days after the trial.	PI (standard 4-point scale) PR (standard 5-point scale) 50% PR? (y/n) Global rating (5-point scale) by patient VAS PI: 'no pain' to 'worst pain imaginable'	Ibuprofen, 400 mg (n = 81); placebo (n = 39).	Ibuprofen significantly superior to placebo for all measures after 30 minutes. 6-hour TOTPAR: ibuprofen 10.9; placebo 2.9.	Patients allowed to remedicate after 2 hours. After remedication PR = 0 and PI = baseline or last score (whichever was greater) for all further time points.	201 analysed. Exclusions: 1 took medication twice, 5 had insufficient pain.	None serious reported; 1 patient in ibuprofen group withdrew as a result of vomiting which investigators did not attribute to medication. Numbers reporting adverse effects (number of effects): ibuprofen 8/81 (13); placebo 1/39 (1).
Gay, *et al.*, 1996	Third molar extraction n = 206 Age range: 18–60 years	RCT, double-blind, single oral dose, parallel groups. 12-hour washout prior to start. Local anaesthetic. Self-assessed at 'regular intervals' for 6 hours. Medication taken when baseline PI moderate to severe.	PI (standard 4-point scale) PR (standard 5-point scale) 50% PR? (y/n) Global rating (5-point scale) by patient VAS PI: 'no pain' to 'worst pain imaginable'	Ibuprofen, 400 mg (n = 41); placebo (n = 39).	Ibuprofen significantly superior to placebo for all summary measures of analgesia ($p < 0.05$). 6-hour TOTPAR: ibuprofen 13.6; placebo 5.2.	Patients allowed to remedicate after 1 hour. If they remedicated earlier data was excluded from efficacy analysis. After remedication PR = 0 and PI = baseline or last score (whichever was greater) for all further time points.	194 analysed. Exclusions: 2 remedicated before 1 hour, 10 failed to complete assess ment within 15 minutes of scheduled time.	None serious reported; no patient withdrew as a result. Numbers reporting adverse effects (number of effects): ibuprofen 3/41 (3); placebo 4/41 (7).
Heidrich, *et al.*, 1985	Orthopaedic surgery n = 120 Age range: 18–65 years	RCT, double-blind, single oral dose, parallel groups. 4-hour washout prior to start. Assessed by trained nurse observer at 0, 0.5 and 1 hour, then hourly for 6 hours. Medication taken when baseline PI moderate to severe.	PI (4-point scale) PR (5-point scale) VAS PI: 'no relief' to 'complete relief' VAS PI: 'no pain' to 'worst pain imaginable' McGill pain questionnaire	Ibuprofen, 400 mg (n = 40); placebo (n = 40).	Orthogonal analyses of variance showed ibuprofen produced greater PR than placebo. 6-hour VAS TOTPAR: ibuprofen 234; placebo 104.	No information given on patients who remedicated.	No information given on any exclusions.	"There were no differences among treatments in terms of side effects. No patient withdrew because of adverse events."

continued

TABLE 28 contd *Studies of ibuprofen versus placebo*

Study	Condition and number of patients	Design, study duration and follow-up	Outcome measures	Dosing regimen	Analgesic outcome results	Remedication	Withdrawals and exclusions	Adverse effects
Hersh, *et al.*, 1993	Third molar extraction n = 254 Age: 16+ years	RCT, double-blind, single oral dose, parallel groups. 4-hour washout prior to start. Local anaesthetic. Self-assessed at 0, 0.5, and 1 hour then hourly for 8 hours (for the first 2 hours in clinic). Medication taken when baseline PI moderate to severe.	PI (standard 4-point scale) PR (standard 5-point scale) Global rating (5-point scale) by patient	Ibuprofen, 400 mg (n = 49); ibuprofen, 200 mg (n = 51); placebo (n = 51).	Ibuprofen at both doses significantly superior to placebo for all measures of analgesia ($p < 0.05$). 6-hour TOTPAR (calculated from the graph): ibuprofen, 400 mg, 9.1; ibuprofen, 200 mg, 6.68; placebo 1.18.	Patients allowed to remedicate after 1 hour. If they remedicated earlier data excluded from efficacy analysis. After remedication PR = 0 and PI = baseline or last score (whichever was greater) for all further time points.	254 analysed. No exclusions.	None serious reported; no patient withdrew as a result. Numbers reporting adverse effects (number of effects): ibuprofen, 400 mg, 6/49 (7); ibuprofen, 200 mg, 4/51 (4); placebo 9/51 (9).
Hersh, *et al.*, 1993	Third molar extraction n = 114 Age: not specified	RCT, double-blind, pre-surgery placebo then single oral dose, parallel groups. Local anaesthetic. Self-assessed at 0, 0.5, and 1 hours, then hourly for 6 hours. Medication taken when baseline pain was of moderate to severe intensity.	PI (standard 4-point scale) PR (standard 5-point scale) 50% PR? (y/n) Global rating (5-point scale) by patient	Placebo then ibuprofen, 400 mg (n = 12); placebo then placebo (n = 16).	Ibuprofen significantly superior to placebo for all summary measures of analgesia. 6-hour TOTPAR: ibuprofen 15.67; placebo 9.00.	Patients allowed to remedicate after 1 hour. After remedication PR = 0 and PI = baseline or last score (whichever was greater) for all further time points.	81 analysed. Exclusions: 19 lost to follow-up, 11 did not require medication, 3 excluded for various protocol violations.	No information given.
Jain, *et al.*, 1986	Third molar extraction n = 260 Age range: 18–65 years	RCT, double-blind, single oral dose, parallel groups. Self-assessed at home at 0 and 1 hours, then hourly for 6 hours. Medication taken when baseline PI moderate to severe.	PI (4-point scale) – standard wording but scale 1–4 PR (non-standard 5-point scale) Global rating (5-point scale) by patient VAS PI: 'no pain' to 'worst ever pain'	Ibuprofen, 400 mg (n = 49); ibuprofen, 200 mg (n = 47); ibuprofen, 100 mg (n = 39); placebo (n = 47).	All ibuprofen doses significantly superior to placebo ($p < 0.001$). 6-hour SPID: ibuprofen, 400 mg, 3.0; ibuprofen, 200 mg, 2.26; ibuprofen, 100 mg, 1.54; placebo –1.73.	Patients allowed to remedicate after 1 hour. If they remedicated earlier data excluded from efficacy analysis. After remedication PR = 0 and PI = last score for all further time points.	227 analysed. Exclusions: 10 remedicated before 1 hour, 19 did not take medication or were lost to follow-up, 2 had mild baseline pain, 1 missed > 2 evaluations, 1 used confounding drugs.	None serious reported; no patient withdrew as a result. Numbers reporting adverse effects (number of effects): ibuprofen, 400 mg, 10/? (12); ibuprofen, 200 mg, 6/? (8); ibuprofen, 100 mg, 13/? (15); placebo 12/? (15).
Jain, *et al.*, 1988	Episiotomy n = 161 Age: 18+ years	RCT, double-blind, single oral dose, parallel groups. 4-hour washout prior to start. Assessed by trained nurse observer at 0, 0.5, and 1 hour, then hourly for 6 hours. Medication taken when baseline pain was of moderate to severe intensity.	PI (standard 4-point scale) PR (standard 5-point scale) Time to meaningful relief Global rating (5-point scale) by patient Overall improvement (7-point scale) by patient	Ibuprofen, 400 mg (n = 49); placebo (n = 48).	Ibuprofen significantly superior to placebo for most summary measures of analgesia ($p < 0.01$). 6-hour TOTPAR: ibuprofen 14.4; placebo 8.61.	Patients allowed to remedicate after 2 hours. If they remedicated earlier data excluded from efficacy analysis. After remedication PR = 0 and PI = baseline or last score (whichever was greater) for all further time points.	147 analysed. Exclusions: 11 remedicated before 2 hours, 2 received confounding agents, 1 was aged under 18 years.	None serious reported; no patient withdrew as a result. Numbers reporting adverse effects (number of effects): ibuprofen 2/49 (2); placebo 1/48 (1).
Kiersch, *et al.*, 1993	Third molar extraction n = 205 Age: 15+ years	RCT, double-blind, single oral dose, parallel groups. 72-hour washout prior to start. Self-assessed at home at 0, 20, 30, 40, 60 minutes, then hourly for 12 hours. Medication taken when baseline PI at least moderate. Review 1/2 days after the trial.	PI (standard 4-point scale) PR (standard 5-point scale) 50% PR? (y/n) Global rating (5-point scale) by patient VAS PI: 'no pain' to 'worst pain imaginable'	Ibuprofen, 200 mg (n = 81); placebo (n = 42).	Ibuprofen significantly superior to placebo for all summary measures of analgesia ($p < 0.001$). 6-hour TOTPAR: ibuprofen 10.3; placebo 3.7.	Patients allowed to remedicate after 2 hours. No information given on how data were then handled.	203 analysed. Exclusions: 2 for protocol violations.	None serious reported; no patient withdrew as a result. Numbers reporting adverse effects (number of effects): ibuprofen 16/81 (20); placebo 5/43 (5).

continued

TABLE 28 contd *Studies of ibuprofen versus placebo*

Study	Condition and number of patients	Design, study duration and follow-up	Outcome measures	Dosing regimen	Analgesic outcome results	Remedication	Withdrawals and exclusions	Adverse effects
Laska, *et al.*, 1986	Third molar extraction n = 200 Age: 16+ years	RCT, double-blind, single oral dose, parallel groups. 4-hour washout prior to start. Self-assessed at 0, 0.5, and 1 hour then hourly for 6 hours. Medication taken when baseline PI moderate to severe.	PI (standard 4-point scale) PR (standard 5-point scale) Global rating by patient Blood serum levels	Ibuprofen, 400 mg (n = ?); ibuprofen, 600 mg (n = ?); ibuprofen, 800 mg (n = ?); placebo (n = ?). Assumed equal distribution of 40 patients per group.	All three doses of ibuprofen significantly superior to placebo for % SPID. 6-hour SPID (calculated from the PI difference graph): ibuprofen, 400 mg, 13.4; ibuprofen, 600 mg, 14.1; ibuprofen, 800 mg, 13.9; placebo 5.3.	Patients allowed to remedicate after 1 hour. If they remedicated earlier data excluded from efficacy analysis. After remedication PR = 0 and PI = baseline or last score (whichever was greater) for all further time points.	195 analysed. Exclusions: 4 remedicated before 1 hour, 1 vomited within 5 minutes of taking study medication.	No withdrawals as a result. Numbers reporting adverse effects (number of effects): ibuprofen, 400 mg, 1/? (1); placebo 3/? (3).
Lavenziana, *et al.*, 1996	Postoperative (inguinal hernia) n = 125 Age range: 18–75 years	RCT, double-blind, single oral dose, parallel groups. 6-hour washout prior to start. Assessed in hospital at 0, 15, 30, 45, 60, 90 minutes, 2 hours then hourly for up to 6 hours. Medication taken when baseline PI moderate to severe.	VAS PI: 'no pain' to 'unbearable pain' Global rating (5-point scale) by patient	Ibuprofen arginine soluble, 400 mg (n = 42); placebo (n = 41).	Patients with 61–80 mm baseline pain significantly superior to placebo ($p < 0.05$). Patients with > 81 mm baseline pain, NSD. VAS SPID from graph: ibuprofen 250; placebo 215.	Patients allowed to remedicate after 1 hour. Patients asked to wait until pain returned to baseline intensity before remedicating but no information given on how data were then handled.	124 analysed. Exclusions: 1 patient for insufficient pain.	None reported.
McQuay, *et al.*, 1996	Third molar extraction n = 218 Age range: 16–53 years	RCT, double-blind, single oral dose, parallel groups. 12-hour washout prior to start. Local anaesthetic. Self-assessed for 6 hours (time points not identified). Medication taken if baseline PI moderate to severe within 2 hours of surgery.	PI (standard 4-point scale) PR (standard 5-point scale) Global rating (5-point scale) by patient VAS PI: 'no pain' – 'worst pain imaginable' VAS PR: 'no relief' to 'complete relief' Random 8-word scale Mood (VAS) Stopwatch to meaningful relief	Ibuprofen, 400 mg (n = 30); ibuprofen, 200 mg (n = 31); placebo (n = 11).	Ibuprofen at both doses significantly superior to placebo for all measures of analgesia. 6-hour TOTPAR (calculated from the graph): ibuprofen, 400 mg, 9.1; ibuprofen, 200 mg, 6.68; placebo 1.18.	Patients allowed to remedicate after 45 minutes If they remedicated earlier data excluded from efficacy analysis. After remedication PR = 0 and PI = baseline for all further time points.	161 analysed. Exclusions: 15 no pain, 10 concurrent illness, 7 analgesics within 48 hours, 4 withdrew before study began, 4 did not attend, 3 previous NSAID allergy, 1 possible pregnancy, 1 migraine after surgery, 1 surgery cancelled, 3 remedicated before 45 minutes.	None serious reported; no patient withdrew as a result. Numbers reporting adverse effects (number of effects): ibuprofen, 400 mg, 2/30 (3); ibuprofen, 200 mg, 4/31 (4); placebo 1/11 (1).
Mehlisch, *et al.*, 1990	Various oral surgery procedures n = 706 Age range: 18–64 years	RCT, double-blind, single oral dose, parallel groups. 6-hour washout prior to start. Self-assessed at 0, 0.5, and 1 hour, then hourly for up to 6 hours. Medication taken when baseline PI moderate to severe.	PI (4-point scale): standard wording scale 1–4 PR (non-standard 4-point scale)	Ibuprofen, 400 mg (n = 306); placebo (n = 85).	Ibuprofen significantly superior for most summary measures of efficacy. 6-hour SPID: ibuprofen 5.84; placebo 0.99.	Patients allowed to remedicate. After remedication PR = 0 and PI = baseline score for all further time points.	697 analysed. Exclusions: 4 lost to follow-up, 4 entered in trial twice (only first entry analysed for efficacy but both included in safety analysis), 1 failed to meet inclusion criteria.	None serious reported; no patient withdrew as a result. Numbers reporting adverse effects (number of effects): ibuprofen 31/310 (?); placebo 12/85 (?).
Mehlisch, *et al.*, 1995	Third molar extraction n = 205 Age: 15+ years	RCT, double-blind, single oral dose, parallel groups. 12-hour washout prior to start. Local anaesthetic. Self-assessed at 0, 15, 30, 45, 60, 90 minutes, 2 hours, then hourly for 6 hours. Medication taken when baseline PI moderate to severe.	PI (standard 4-point scale) PR (standard 5-point scale) Global rating (5-point scale) by patient	Ibuprofen, 400 mg (n = 98); placebo (n = 40).	Ibuprofen significantly superior to placebo for all measures of analgesia ($p < 0.05$). 6-hour TOTPAR: ibuprofen 14.39; placebo 2.62.	Patients allowed to remedicate after 1 hour (but encouraged to wait for 4 hours). If remedicated earlier data excluded from efficacy analysis. After remedication PR and PID = 0 for all further time points.	239 analysed. Exclusion: 1 patient had only 1 molar removed and failed to complete diary.	None serious reported; no patient withdrew as a result. Numbers reporting adverse effects (number of effects): ibuprofen 12/98 (?); placebo 4/40 (?).

continued

TABLE 28 contd *Studies of ibuprofen versus placebo*

Study	Condition and number of patients	Design, study duration and follow-up	Outcome measures	Dosing regimen	Analgesic outcome results	Remedication	Withdrawals and exclusions	Adverse effects
Nelson, *et al.*, 1994	Third molar extraction n = 183 Age: 15+ years	RCT, double-blind, single oral dose, parallel groups. 12-hour washout prior to start. Local anaesthetic. Self-assessed at 0, 15, 30, 45, 60, 90 minutes, 2 hours, then hourly for 6 hours. Medication taken when baseline PI moderate to severe.	PI (standard 4-point scale) PR (standard 5-point scale) 50% PR? (y/n) Global rating (5-point scale) by patient	Ibuprofen, 200 mg (n = 75); placebo (n = 40).	Ibuprofen significantly superior to placebo for PR and PI differences from 30 minutes onwards. 6-hour TOTPAR: ibuprofen 12.31; placebo 5.56.	Patients allowed to remedicate after 1 hour. If they remedicated earlier data excluded from efficacy analysis. After remedication PR = 0 and PI = 0 for all further time points.	180 analysed. Exclusions: 2 remedicated before 1 hour, 1 did not record baseline pain intensity.	None serious reported; no patient withdrew as a result. Numbers reporting adverse effects (number of effects): ibuprofen 16/77 (?); placebo 11/41 (?).
Pagnoni, *et al.*, 1996	Caesarean section n = 92 Age: 18+ years	RCT, double-blind, single oral dose, parallel groups. 6-hour washout prior to start. General anaesthetic. Assessed in hospital at 0, 15, 30, 45, 60, 90 minutes, 2 hours, then hourly for 6 hours. Medication taken when baseline pain was > 55 mm.	VAS PI: 'no pain' to 'unbearablepain' Global rating (5-point scale) by patient	Ibuprofen arginine soluble, 400 mg (n = 30); placebo (n = 32).	Sum of PID and mean AUC showed ibuprofen significantly superior to placebo ($p < 0.001$); mean peak PI difference value was also signifi-cantlysuperior to placebo ($p < 0.05$). 6-hour VAS SPID from graph: ibuprofen 181 mm; placebo 65 mm.	Patients allowed to remedicate after 1 hour. If they remedicated earlier data excluded from efficacy analysis. After remedication PI = last recorded value for all further time points.	92 analysed. No exclusions.	None reported.
Parker, *et al.*, 1986	Tonsillectomy n = 139 Age range: 16–66 years	RCT, double-blind, single oral dose then multiple doses, parallel groups. General anaesthetic. Assessed in hospital at 0, 0.5, 1 hours, then hourly for 4 hours. Medication taken when baseline PI moderate to severe.	PI (non-standard 9-point scale) PR (standard 5-point scale) 50% PR? (y/n) Observers noted their impression of patients' progress	Ibuprofen syrup, 600 mg (n = 44); placebo (n = 33).	Ibuprofen significantly superior to placebo at 30 minutes and 1 hour. 4-hour TOTPAR: ibuprofen 10.92; placebo 9.37.	No information given on patients who remedicated.	110 analysed. No information given on 29 exclusions.	No details given of those occurring during single dose. For multiple doses, 1 patient in placebo group withdrew as a result. Numbers reported were similar for both groups.
Schachtel, *et al.*, 1989	Episiotomy n = 115 Age range: 16–37 years	RCT, double-blind, single oral dose then multiple doses, parallel groups. 4-hour washout prior to start. Assessed in hospital at 0, 0.5, 1 hours, then hourly for 4 hours. Medication taken when baseline PI moderate to severe.	PI (standard 4-point scale) PR (standard 5-point scale) Global rating (5-point scale) by patient	Ibuprofen, 400 mg (n = 36); placebo (n = 38).	Ibuprofen significantly superior to placebo for all measures of analgesia ($p < 0.05$) at least. 4-hour TOTPAR: ibuprofen 10.4; placebo 5.5.	Patients allowed to remedicate after 1 hour. If they re-medicated earlier data excluded from efficacy analysis. After remedication PR = 0 and PI = last or baseline (which ever was greater) for all further time points.	111 analysed. Exclusions: 4 remedicated before 1 hour.	None reported.
Seymour, *et al.*, 1991	Third molar extraction n = 205 Age: Adult	RCT, double-blind, single oral dose, parallel groups. General anaesthetic. Assessed in hospital by same observer at 0, 10, 20, 30, 45, 60, 90 minutes, 2 hours, then hourly for 6 hours. Medication taken when baseline pain was > 30 mm.	VAS PI: 'no pain' to 'unbear-able pain' Global rating (5-point scale) by patient	Study 1: ibu-profen (tablets), 400 mg (n = 31); ibuprofen (liquid in gelatin capsules), 400 mg (n = 32); placebo (n = 32). Study 2: ibuprofen (tablets), 400 mg (n = 30); ibuprofen (soluble), 400 mg (n = 32); placebo (n = 30).	Study 1: both forms of ibuprofen significantly superior to placebo; no significant difference between 2 active groups. Study 2: soluble ibu-profen significantly superior to placebo from 20 minutes; tablets from 30 minutes. 6-hour VAS SPID: gelatin 233 mm; tablets 243 mm; placebo 120 mm; soluble 228 mm; tablets 214 mm; placebo 86 mm.	Patients allowed to remedicate. After remedication PI = last score for all further time points.	187 analysed. Claimed to have enrolled only 180?	None serious reported; no patient withdrew as a result. Only 1 patient from the placebo group (Study 1) reported an adverse event.

continued

TABLE 28 contd *Studies of ibuprofen versus placebo*

Study	Condition and number of patients	Design, study duration and follow-up	Outcome measures	Dosing regimen	Analgesic outcome results	Remedication	Withdrawals and exclusions	Adverse effects
Seymour, *et al.*, 1996	Third molar extraction n = 148 Age: Adult	RCT, double-blind, single oral dose, parallel groups. General anaesthetic. Assessed in hospital by nurse observer at 0, 10, 20, 30, 45, 60, 75, 90,120, 150 minutes, 3 hours, then hourly for 6 hours. Medication taken when baseline pain was > 30 mm.	PI (VAS): 'no pain' to 'unbearable pain' Global rating (5-point scale) by patient	Ibuprofen (tablets), 600 mg (n = 17); ibuprofen (soluble), 600 mg (n = 17); ibuprofen (tablets), 400 mg (n = 15); ibuprofen (soluble), 400 mg (n = 16); ibuprofen (tablets), 200 mg (n = 18); ibuprofen (soluble), 200 mg (n = 17); placebo (n = 19).	All ibuprofen treatments except 200 mg resulted in significantly less pain than placebo for all efficacy measures ($p < 0.05$). 6-hour VAS SPID: ibuprofen, 600 mg tablets, 230; ibuprofen, 600 mg soluble, 148; ibuprofen, 400 mg tablets, 258; ibuprofen, 400 mg soluble, 238; ibuprofen, 200 mg tablets, 140; ibuprofen, 200 mg soluble, 198; placebo 44.	Patients allowed to remedicate. After remedication PI = last score for all further time points.	199 analysed. Exclusions: 4 for 'unwanted effects', 25 failed to reach sufficient baseline PI.	Reported by 4 patients: 3 took ibuprofen (dose not specified) and 1 placebo.
Sunshine, *et al.*, 1983	Episiotomy n = 115 Age: 18+ years	RCT, double-blind, single oral dose, parallel groups. 4-hour washout prior to start. Assessed in hospital by same observer at 0, 0.5, 1 hours, then hourly for 4 hours. Medication taken when baseline PI moderate to severe.	PI (standard 4-point scale) PR (non-standard 5-point scale) (percentages not descriptive wording) Global rating of medication (4-point scale) by patient Global rating of personal improvement (7-point scale) by patient	Ibuprofen, 400 mg (n = 30); placebo (n = 30).	Ibuprofen significantly superior to placebo for all measures of analgesia from 1 hour onwards. 4-hour SPID: ibuprofen 6.47; placebo 1.12.	Patients allowed to remedicate after 1 hour. If they remedicated earlier data excluded from efficacy analysis. After remedication PR = 0 and PI = last for all further time points.	120 analysed. No exclusions.	None reported.
Sunshine, *et al.*, 1987	Episiotomy, Caesarean section or gynaecological surgery n = 200 Age: not specified	RCT, double-blind, single oral dose, parallel groups. 4-hour washout prior to start. Assessed in hospital by same observer at 0, 0.5, 1 hours, then hourly for 4 hours. Medication taken when baseline PI moderate to severe.	PI (standard 4-point scale) PR (non-standard 5-point scale) (percentages not descriptive wording) Global rating of medication (4-point scale) by patient Global rating of personal improvement (7-point scale) by patient	Ibuprofen, 400 mg (n = 38); placebo (n = 40).	All active treatments significantly superior to placebo for TOTPAR and all except codeine for SPID. 4-hour SPID: ibuprofen 8.1; placebo 5.2.	Patients allowed to remedicate after 1 hour. If they remedicated earlier data excluded from efficacy analysis. After remedication PR = 0 and PI = last for all further time points.	195 analysed. Exclusions: 1 non-compliant with washout period, 4 did not complete evaluations.	None reported in either placebo or ibuprofen groups.

Each point in *Figure 28* represents one trial with the proportion of patients achieving at least 50% pain relief on the study drug plotted on the *y* axis, and the proportion of patients achieving the same end point with placebo on the *x* axis.

In one trial,[19] a syrup formulation of ibuprofen was used, in two trials[20,21] soluble ibuprofen and liquid in gelatin capsules were used, in two trials[22,23] ibuprofen lysine was used, and in two trials[24,25] soluble ibuprofen arginine was used. When the results from these more readily absorbed formulations were pooled and compared with those for the standard tablet formulation, no differences were found in the relative benefit or NNTs. The NNT for a single dose of ibuprofen, 400 mg, standard formulation tablets (1356 patients) compared with placebo was 2.8 (95% CI, 2.5–3.1) and for ibuprofen, 400 mg, soluble formulations (250 patients) the NNT was 2.5 (95% CI, 2.1–3.1). All formulations were therefore pooled for the overall analysis.

The single data set for ibuprofen, 50 mg, showed no significant difference from placebo. The pooled relative benefits for ibuprofen, 100 mg, 200 mg, 400 mg and 600 mg, were significantly different from placebo, as was the single data set for ibuprofen, 800 mg (*Table 29*). At a dose of 50 mg, ibuprofen had an NNT of 3.6 (95% CI, 2.5–6.1) for at least

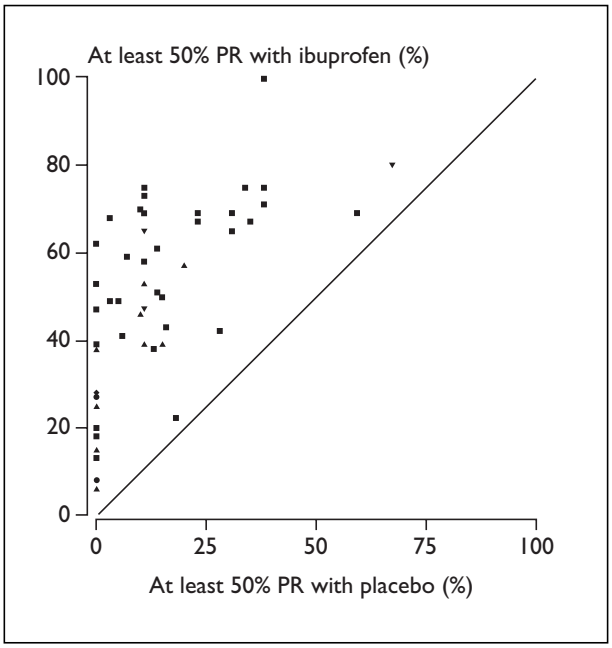


FIGURE 28 *Ibuprofen against placebo in postoperative pain (◆, ibuprofen, 50 mg; ●, ibuprofen, 100 mg; ▲, ibuprofen, 200 mg; ■, ibuprofen, 400 mg; ▼, ibuprofen, 600 mg)*

50% pain relief over 4–6 hours compared with placebo in pain of moderate to severe intensity; at 100 mg, the NNT was 5.6 (95% CI, 3.8–9.9), at 200 mg, the NNT was 3.3 (95% CI, 2.8–4.0), at 400 mg, the NNT was 2.7 (2.5–3.0), at 600 mg, the NNT was 2.4 (1.9–3.3) and, at 800 mg, the NNT was 1.6 (1.3–2.2). The dose response for ibuprofen is presented in *Figure 29*.

When a fixed CER of 19% was applied, there was a clear dose response with no overlap in CIs except for the 600 mg and 800 mg doses *(Table 29;* see also *Figure 32)*.

Drug-related study withdrawals occurred rarely. In one study,[26] one withdrawal on ibuprofen was for vomiting which the authors did not attribute to the

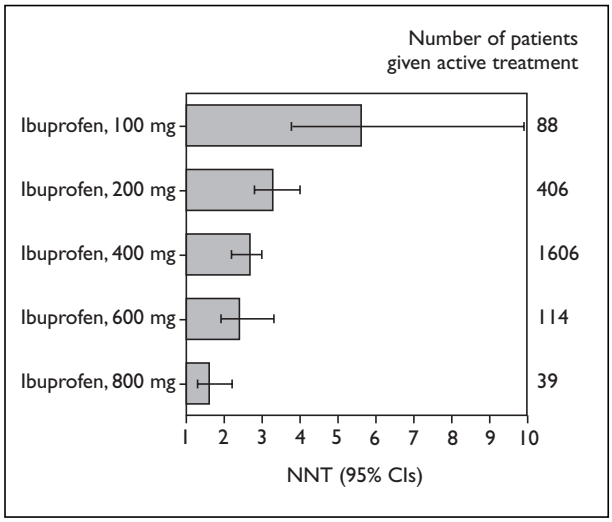


FIGURE 29 *Dose response for ibuprofen trials*

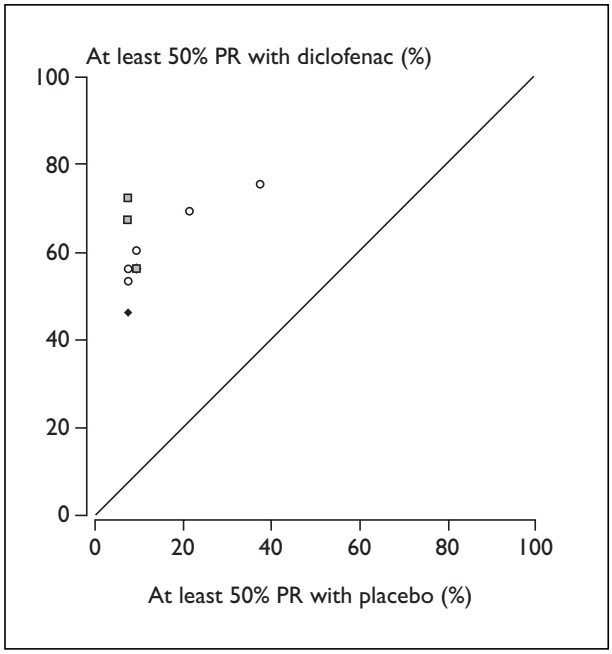


FIGURE 30 *Diclofenac against placebo in postoperative pain (◆, diclofenac, 25 mg; ○, diclofenac, 50 mg; □, diclofenac, 100 mg)*

TABLE 29 *Summary of relative benefit and NNTs for trials of ibuprofen versus placebo*

Number of trials	Dose of ibuprofen (mg)	Number of patients with > 50% PR: ibuprofen	Number of patients with > 50% PR: placebo	RB – random effects model (95% CI)	NNT (95% CI)	NNT with 19% CER (95% CI)
1	50	16/57	0/51	144 (0.3– > 1000)	3.6 (2.5–6.1)	12.5 (4.1–∞)
2	100	16/88	0/98	72 (16–318)	5.6 (3.8–9.9)	–100 (10–∞)
8	200	151/406	22/320	3.5 (2.3–5.3)	3.3 (2.8–4.0)	5.6 (4.1–8.5)
30	400	858/1606	214/1292	3.3 (2.5–4.3)	2.7 (2.5–3.0)	2.9 (2.7–3.2)
3	600	90/114	40/108	2.5 (1.2–5.5)	2.4 (1.9–3.3)	1.7 (1.4–2)
1	800	39/39	14/37	2.6 (1.8–4)	1.6 (1.3–2.2)	1.2 (1.1–1.5)

medication. In one study,[21] there were three withdrawals on ibuprofen and one on placebo for vomiting soon after ingestion of the study drug. In another study,[19] one patient withdrew on placebo. The studies reported a variable incidence of minor adverse events which were all mild and transient, with no difference in incidence between ibuprofen and placebo.

Diclofenac versus placebo

Although the search identified nearly 2000 trials, the majority were in chronic pain or the drug was administered before the patient experienced pain. Over 500 reports were found of trials involving rectal, intravenous and intramuscular diclofenac. Predominantly, the reports were not in established postoperative pain, were not placebo-controlled or did not use standard pain outcome measures. Only six trials fulfilled the inclusion criteria; all were for oral diclofenac (528 patients were given diclofenac and 312 placebo). Five reports identified by the search could not be obtained despite attempts to contact the authors, ordering through the British Library and help from the librarians at Novartis and Knoll Pharmaceuticals.[27–31] Details of the six included studies are presented in *Table 30,* with their references being given at the end of this chapter. The authors of one study were contacted for information regarding the number of patients in each treatment arm;[32] however, they were unable to provide this information and so an equal split of 50 patients per group was assumed.

Five trials investigated oral surgery pain (third molar extraction with bone removal) and one pain following gynaecological surgery. The doses of diclofenac prescribed were 25 mg in one trial, 50 mg in six trials and 100 mg in three trials. In three trials the immediate release diclofenac potassium formulation was used,[32–34] and in two trials dispersible diclofenac was used.[35,36] In one trial, both the immediate release and enteric-coated formulations were used.[37] To ensure comparability, only data from the immediate release formulation were included.

CER values ranged from 8% to 38% (median 10%) *(Figure 30)*. The EER for the single trial of diclofenac, 25 mg, was 46%. The EER for diclofenac, 50 mg, varied between 53% and 75% (median 58%) and for diclofenac, 100 mg, between 56% and 72% (median 67%). The 100 mg data set was homogeneous but the 50 mg data set was not.

The pooled relative benefits for all doses of diclofenac versus placebo were significant *(Table 31)*. At a dose of 25 mg, diclofenac had an NNT of 2.6 (95% CI, 1.9–4.5) for at least 50% pain relief over 4–6 hours compared with placebo in pain of moderate to severe intensity. The NNT at 50 mg was 2.3 (95% CI, 2.0–2.7) and at 100 mg 1.8 (95% CI, 1.5– 2.1), with overlapping CIs. The dose response for diclofenac is shown in *Figure 31*.

With a fixed value of CER of 19%, the NNT for diclofenac, 25 mg, was 3.9 (95% CI, 2.3–12), for 50 mg 2.3 (95% CI, 2.0–2.7) and for 100 mg 2.2 (95% CI, 1.8–2.8), with overlapping CIs *(Table 29; Figure 32)*.

Drug-related study withdrawals rarely occurred. In one study,[33] there was one withdrawal on diclofenac, 100 mg, for nausea and vomiting. The studies reported a variable incidence of minor adverse events none of which were serious and there was no difference in incidence between diclofenac and placebo.

Diclofenac versus ibuprofen

There were two direct comparisons of diclofenac, 50 mg, and ibuprofen, 400 mg.[35,36] Both trials were in dental pain (third molar removal); 118 patients received diclofenac and 112 patients ibuprofen. There was no significant difference between diclofenac, 50 mg, and ibuprofen, 400 mg (relative benefit 1.0; 95% CI, 0.9–1.2).

Comment

A single dose of ibuprofen, 400 mg, had an NNT of 2.7 for at least 50% pain relief compared with placebo. This means that one from every three patients with pain of moderate to severe intensity will experience at least 50% pain relief with ibuprofen which they would not have had with placebo. The equivalent NNT for a single dose of ibuprofen, 600 mg, was 2.4 and for ibuprofen, 200 mg, 3.3, showing a dose response although the CIs overlapped *(Table 29)*. When a fixed CER was used to smooth out the CER variations of individual trials, the CIs for the NNTs did not overlap, supporting the dose–response finding *(Figure 32)*. Moreover, the use of a fixed (population) CER had little effect on the NNT in circumstances where there were either large numbers of patients or where there were large effects *(Tables 29* and *31)*. Only in small trials with limited analgesic efficacy (low doses) did the use of the fixed CER alter the NNT significantly.

A single dose of diclofenac, 50 mg, had an NNT of 2.3 for at least 50% pain relief compared with

TABLE 30 *Studies of diclofenac versus placebo*

Study	Condition and number of patients	Design, study duration and follow-up	Outcome measures	Dosing regimen	Analgesic outcome results	Remedication	Withdrawals and exclusions	Adverse effects
Ahlstrom, *et al.*, 1993	Third molar extraction n = 127 Age range: 18 – 40 years	RCT, double-blind, single oral dose, parallel groups. 4-hour washout prior to start. Evaluated at 0, 20, 40 minutes, 1 hour then hourly for 6 hours. Medication taken when baseline pain was at least moderate intensity (> 30 mm).	VAS PI: 'no pain at all' to 'agonising pain' Global rating by patient	Diclofenac (drinkable), 50 mg (n = 35); placebo (n = 30).	Diclofenac significantly superior to placebo by 40 minutes (p = 0.001); this continued for 6 hours. TOTPI and SPID: diclofenac significantly superior to placebo (p < 0.0001). 6-hour SPID: diclofenac 173 mm; placebo 32 mm.	Patients allowed to remedicate after 1 hour. After remedication PI = last score was carried forward for all further time points.	97 analysed. Exclusions: 30 for various protocol violations.	None serious reported; no patient withdrew as a result. Numbers reporting adverse effects (number of effects): diclofenac 6/35 (?); placebo 2/30 (?).
Bakshi, *et al.*, 1992	Third molar extraction n = 180 Age: adult	RCT, double-blind, single oral dose, parallel groups. 4-hour washout prior to start. Self-assessed at 0, 15, 30 minutes, 1 hour, then hourly for 6 hours. Medication taken when baseline pain was of at least moderate intensity.	PI (standard 4-point scale) PR (non-standard 4-point scale) 50% PR? (y/n) Global rating (4-point scale) by patient	Diclofenac potassium (sugar-coated), 50 mg (n = 51); diclofenac sodium (enteric-coated), 50 mg (n = 54); placebo (n = 46).	Diclofenac potassium significantly superior to placebo for SPID, TOTPAR, MAXPID and MAXPAR (p < 0.001). Diclofenac sodium significantly superior to placebo for SPID (p = 0.023) and MAXPID (p = 0.018). 6-hour SPID: diclofenac potassium 7.8; diclofenac sodium 5.3; placebo 2.6.	Patients allowed to remedicate after 1 hour. If they remedicated earlier data excluded from efficacy analysis. After remedication PR = 0 and PI = last score or baseline (whichever was greater) for all further time points.	151 analysed. Exclusions: 26 did not require medication, 3 lost to follow-up.	None serious reported; no patient withdrew as a result. Numbers reporting adverse effects (number of effects): diclofenac potassium 3/51 (5); diclofenac sodium 1/54 (1); placebo 3/46 (3).
Bakshi, *et al.*, 1994	Third molar extraction n = 257 Age range: adults up to 65 years	RCT, double-blind, single oral dose, parallel groups. Local anaesthetic. Self-assessed at 0, 20, 40 minutes, 1, 1.5, 2 hours, then hourly for 6 hours. Medication taken when baseline pain was of at least severe intensity.	VAS PI: 'no pain' to 'pain could not be worse' PR (5-point scale): none, poor, moderate, sufficient, total Global rating (non-standard 5-point scale) by patient and by observer	Diclofenac (dispersible), 50 mg (n = 83); placebo (n = 82).	Diclofenac significantly superior to placebo for TOTPAR and both global ratings (p < 0.01). 6-hour TOTPAR: diclofenac 15.45; placebo 8.85.	Patients allowed to remedicate after 1 hour. If they remedicated earlier data excluded from efficacy analysis. After remedication PR = 0 and PI = last score for all further time points.	245 analysed. Exclusions: 9 did not experience severe pain, 2 remedicated before 1 hour, 1 completed diary incorrectly.	None serious reported; no patient withdrew as a result. Numbers reporting adverse effects (number of effects): diclofenac 4/83 (?); placebo 5/82 (?).
Hebertson, *et al.*, 1994	Gynaecological surgery n = 217 Age: 16+ years	RCT, double-blind, single oral dose, parallel groups. 4-hour washout prior to start. Assessed by observer at 0, 0.5, 1 hour, then hourly for 8 hours. Medication taken when baseline pain was of moderate to severe intensity and groups stratified by baseline PI.	PI (standard 4-point scale) PR (standard 5-point scale) Global rating (5-point scale) by patient	Diclofenac, 50 mg (n = 52); diclofenac, 100 mg (n = 52); placebo (n = 52).	Both diclofenac doses significantly superior to placebo for PR at each time point from 1 hour onwards. 6-hour TOTPAR from graph: diclofenac, 50 mg, 12.1; diclofenac, 100 mg, 12.2; placebo 3.7.	Patients allowed to remedicate after 1 hour. If they remedicated earlier data excluded from efficacy analysis. After remedication patients discontinued in study; no information given on how their data were then handled.	209 analysed for at least 1 efficacy analysis. 194 analysed for 8-hour SPID and TOTPAR. No information given on any exclusions.	All were gastrointestinal except one in placebo group (not defined); 1 withdrew from diclofenac, 100 mg, for nausea and vomiting. Numbers reporting adverse effects (number of effects): diclofenac, 50 mg, 3/54 (?); diclofenac, 100 mg, 2/55 (?); placebo 2/54 (?).
Mehlisch, *et al.*, 1994	Third molar extraction n = 208 Age range: 16–70 years	RCT, double-blind, single oral dose, parallel groups. Local anaesthetic. 4-hour washout prior to start. Self-assessed at 0, 0.5, 1 hour, then hourly for up to 8 hours. Medication taken when baseline pain was of moderate to severe intensity.	PI (standard 4-point scale) PR (standard 5-point scale) Global rating (5-point scale) by patient	Diclofenac, 50 mg (n = 53); diclofenac, 100 mg (n = 52); placebo (n = 52).	Both diclofenac groups significantly superior to placebo for all time points with regard to PR. 6-hour TOTPAR from graph: diclofenac, 50 mg, 11.6; diclofenac, 100 mg, 14.3; placebo 3.3.	Patients allowed to remedicate after 2 hours. After remedication PR = 0 and PI = last score for all further time points.	There were no exclusions.	None serious reported; no patient withdrew as a result of adverse events. Numbers reporting adverse effects (number of effects): diclofenac, 50 mg, 2/53 (?); diclofenac, 100 mg, 2/52 (?); placebo 2/52 (?).
Nelson, *et al.*, 1994	Third molar extraction n = 255 Age range: 16–70 years	RCT, double-blind, single oral dose, parallel groups. Local anaesthetic. 4-hour washout prior to start. Self-assessed at 0, 0.5, 1 hour then hourly for up to 8 hours. Medication taken when baseline pain was of moderate to severe intensity.	PI (standard 4-point scale) PR (standard 5-point scale)	Diclofenac, 25 mg (n = ?); diclofenac, 50 mg (n = ?); diclofenac, 100 mg (n = ?); placebo (n = ?). NB: assumed equal distribution of 51 per group (including aspirin group).	All diclofenac groups significantly superior to placebo for PR from 1 hour onwards. 6-hour TOTPAR from graph: diclofenac, 25 mg, 10.5; diclofenac, 50 mg, 12.2; diclofenac, 100 mg, 14.8; placebo 3.6.	Patients allowed to remedicate at 1 hour, no information given as to how their data were then handled.	252 analysed for at least 1 efficacy analysis. No exclusions given other than remedicators.	None serious reported; no patient withdrew as a result of adverse events. Numbers reporting adverse effects (number of effects): diclofenac, 25 mg, 5 (?); diclofenac, 50 mg, 4 (?); diclofenac, 100 mg, 4 (?); placebo 6 (?).

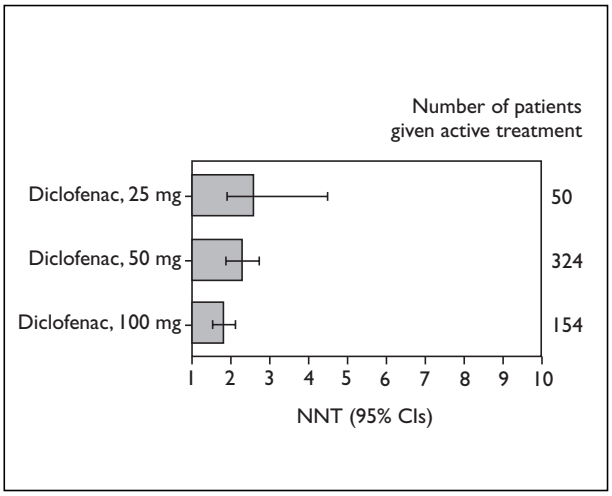


FIGURE 31 *Dose response for diclofenac trials*

placebo. The equivalent NNT for diclofenac, 100 mg, was 1.8 and for diclofenac, 25 mg, 2.6, indicating a dose response although the CIs overlapped *(Table 31)*. With a fixed CER of 19%, the CIs for the two higher doses overlapped completely *(Table 31; Figure 32)*.

Diclofenac is widely regarded as a more effective NSAID than ibuprofen. These results do not support this conclusion. When diclofenac, 50 mg, and ibuprofen, 400 mg, were compared directly there was no significant difference between them. When compared with placebo, diclofenac, 50 mg, and ibuprofen, 600 mg, had very similar NNTs with complete overlap of CIs. Single trials of NSAIDs have often reported flat dose–response curves, typified by that for diclofenac presented in

TABLE 31 *Summary of relative benefit and NNTs for trials of diclofenac versus placebo*

Number of trials	Dose of diclofenac (mg)	Number of patients with > 50% PR: diclofenac	Number of patients with > 50% PR: placebo	RB – random effects model (95% CI)	NNT (95% CI)	NNT with 19% CER (95% CI)
1	25	23/50	4/50	5.8 (2.1–15.4)	2.6 (1.9–4.5)	3.9 (2.3–12.1)
6	50	203/324	57/312	4.3 (2.4–7.8)	2.3 (2–2.7)	2.3 (2–2.7)
3	100	100/154	13/154	7.2 (5.5–9.4)	1.8 (1.5–2.1)	2.2 (1.8–2.8)

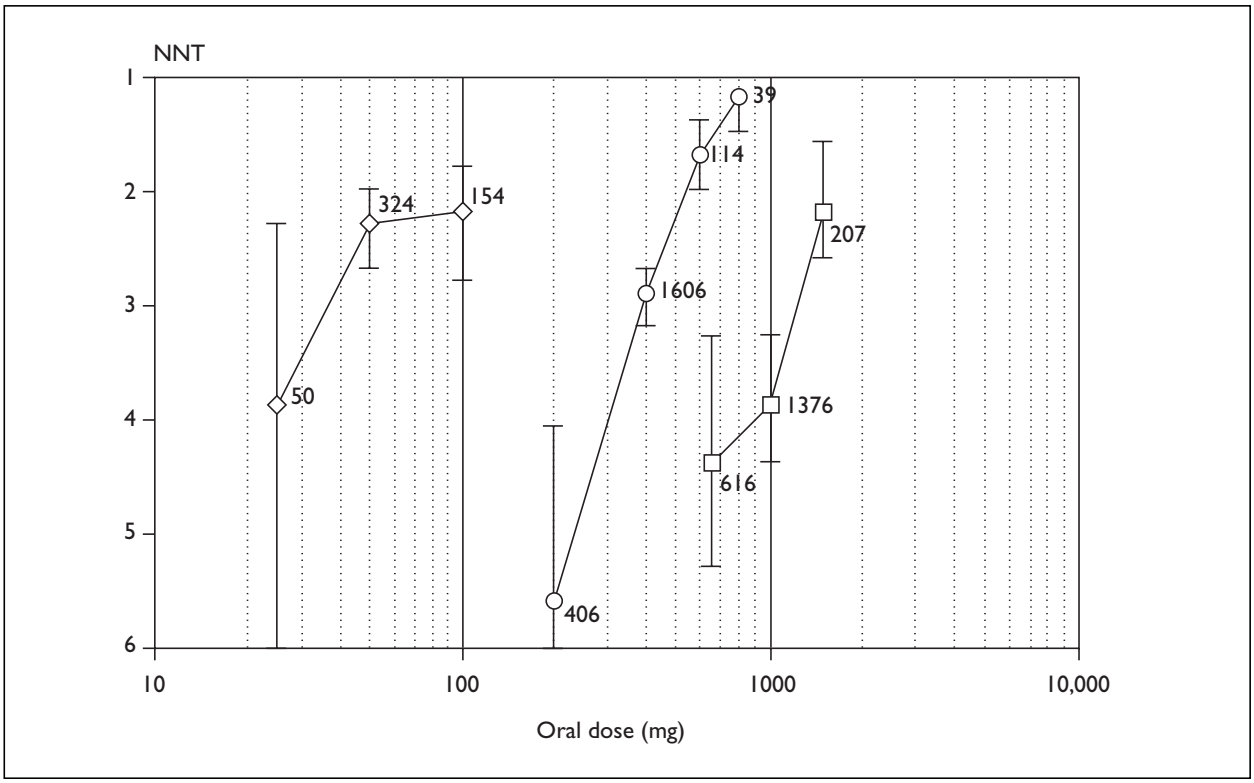


FIGURE 32 *Dose response for diclofenac (◇), ibuprofen (○) and paracetamol (□) using a fixed 19% placebo response rate. Numbers indicate patients given active treatments*

Figure 32. With the advantage of the much larger numbers of patients in this meta-analysis, the 'true' dose response for ibuprofen is shown *(Figure 32),* with the interesting finding that higher doses (> 1 g) of paracetamol also follow traditional dose–response curve contours.

The issue of the relative efficacy of the two drugs therefore comes down to dose *(Figures 29, 31* and *32)* and safety.[38] Ibuprofen, 400 mg, is one-sixth of the maximum daily dose. Diclofenac, 50 mg, is one-third of the maximum daily dose. This may explain prescriber confusion.

References

1. Collins SL, Moore A, McQuay HJ, Wiffen PJ. Oral ibuprofen and diclofenac in postoperative pain: a quantitative systematic review. *Eur J Pain*. In press.
2. Government Statistical Service. Prescription cost analysis for England 1996. London: Department of Health, 1997.
3. Jadad AR, Carroll D, Moore A, McQuay HJ. Developing a database of published reports of randomised clinical trials in pain research. *Pain* 1996;**66**:239–46.
4. Reynolds JEF. Martindale: the extra pharmacopoeia. 30th edition. London: Pharmaceutical Press, 1993:292–314.
5. Collins SL, Moore RA, McQuay HJ. The visual analogue pain intensity scale: what is moderate pain in millimetres? *Pain* 1997;**72**:95–7.
6. L'Abbé KA, Detsky AS, O'Rourke K. Meta-analysis in clinical research. *Ann Intern Med* 1987;**107**:224–33.
7. Cooper SA. Single-dose analgesic studies: the upside and downside of assay sensitivity. In: Max MB, Portenoy RK, Laska EM, editors. The design of analgesic clinical trials. *Adv Pain Res Ther* 1991;**18**:117–24.
8. Moore A, McQuay H, Gavaghan D. Deriving dichotomous outcome measures from continuous data in randomised controlled trials of analgesics. *Pain* 1996;**66**:229–37.
9. Moore A, McQuay H, Gavaghan D. Deriving dichotomous outcome measures from continuous data in randomised controlled trials of analgesics: verification from independent data. *Pain* 1997;**69**:127–30.
10. Moore A, Moore O, McQuay H, Gavaghan D. Deriving dichotomous outcome measures from continuous data in randomised controlled trials of analgesics: use of pain intensity and visual analogue scales. *Pain* 1997;**69**:311–15.
11. DerSimonian R, Laird N. Meta-analysis of clinical trials. *Control Clin Trials* 1986;**7**:177–88.
12. Cook RJ, Sackett DL. The number needed to treat: a clinically useful measure of treatment effect. *BMJ* 1995;**310**:452–4.
13. Moore RA, Collins S, McQuay HJ. Variation in the placebo effect; the impact on individual trials and consequences for meta-analysis. In: Jensen TS, Hammond DL, Jensen TS, editors. Proceedings 8th World Congress on Pain. Progress in pain research and management, vol. 2. Seattle: IASP Press, 1997.
14. Moore A, Collins S, Carroll D, McQuay H. Paracetamol with and without codeine in acute pain: a quantitative systematic review. *Pain* 1997;**70**:193–201.
15. Moore RA, McQuay HJ. Single-patient data meta-analysis of 3453 postoperative patients: oral tramadol versus placebo, codeine and combination analgesics. *Pain* 1997;**69**:287–94.
16. Collins SL, Edwards J, Moore A, McQuay HJ. Single dose oral dextropropoxyphene in post-operative pain: a quantitative systematic review. *Eur J Clin Pharmacol* 1998;**54**:107–12.
17. Edwards JE, McQuay HJ, Moore RA. Systematic review: dihydrocodeine in postoperative pain. Cochrane Library. In press.
18. Laska EM, Sunshine A, Marrero I, Olson N, Siegel C, McCormick N. The correlation between blood levels of ibuprofen and clinical analgesic response. *Clin Pharmacol Ther* 1986;**40**:1–7.
19. Parker D, Gibbin K, Noyelle R. Syrup formulations for post tonsillectomy analgesia: a double blind study comparing ibuprofen, aspirin and placebo. *J Laryngol Otol* 1986;**100**:1055–60.
20. Seymour RA, Hawkesford JE, Weldon M, Brewster D. An evaluation of different ibuprofen preparations in the control of postoperative pain after third molar surgery. *Br J Clin Pharmacol* 1991;**31**:83–7.
21. Seymour R, WardBooth P, Kelly P. Evaluation of different doses of soluble ibuprofen and ibuprofen tablets in postoperative dental pain. *Br J Oral Maxillofac Surg* 1996;**34**:110–14.
22. Mehlisch DR, Jasper RD, Brown P, Korn SH, McCarroll K, Murakami AA. Comparative study of ibuprofen lysine and acetaminophen in patients with postoperative dental pain. *Clin Ther* 1995;**17**:852–60.
23. Nelson SL, Brahim JS, Korn SH, Greene SS, Suchower LJ. Comparison of single-dose ibuprofen lysine, acetylsalicylic acid, and placebo for moderate-to-severe postoperative dental pain. *Clin Ther* 1994;**16**:458–65.
24. Laveneziana D, Riva A, Bonazzi M, Cipolla M, Migliavacca S. Comparative efficacy of oral ibuprofen arginine and intramuscular ketorolac in patients with postoperative pain. *Clin Drug Invest* 1996;**11**:8–14.

25. Pagnoni B, Ravanelli A, Degradi L, Rossi R, Tiengo M. Clinical efficacy of ibuprofen arginine in the management of postoperative pain associated with suction termination of pregnancy. A double-blind placebo-controlled study. *Clin Drug Invest* 1996;**11**:27–32.

26. Fricke J, Halladay S, Francisco C. Efficacy and safety of naproxen sodium and ibuprofen for pain relief after oral surgery. *Curr Therapeut Res Clin Exp* 1993;**54**:619–27.

27. Carlos D. Comparative study of the efficacy of diclofenac Na, meperidine HCl and nalbuphine HCl in post-operative analgesia. *Philip J Intern Med* 1984;**22**:51–5.

28. Frezza R, Bolognesi P, Bernardi F. Confronto sull'azione di tre antinfiammatori non steroidei nel controllo de dolore odontoiatrico post-chirurgico. Efficaaci i FANS contro il dolore [Comparison of the action of 3 non-steroidal anti-inflammatory agents in the control of post-operative pain. Effectiveness of NSAID against pain]. *Attual Dent* 1985;**1**:40–2.

29. Iqbal K, Biswas G, Mondo S, Afzalunnessa B. Assessment of post operative analgesia: a comparative study of pethidine and diclofenac sodium. *J Bangladesh Coll Physicians Surg* 1986;**4**:1–7.

30. Joubert JJ. An assesment of the efficacy and tolerability of Voltaren in the treatment of inflammation after extraction of teeth. *J Dent Assoc S Africa* 1977;**32**:581–3.

31. Vigneron JR, Thys R. Etude de l'action anti-inflammatoire et antalgique du diclofenac en traumatologie et chirurgie orthopedique [Study of the anti-inflammatory and analgesic actions of diclofenac in traumatology and orthopedic surgery]. *Rev Med Liege* 1977;**32**:10–14.

32. Nelson S, Brahim J. An evaluation of the analgesic efficacy of diclofenac potassium, aspirin, and placebo in postoperative dental pain. *Today's Ther Trends* 1994;**12**:3–14.

33. Hebertson R, Storey N. The comparative efficacy of diclofenac potassium, aspirin, and placebo in the treatment of patients with pain following gynecologic surgery. *Today's Ther Trends* 1994;**12**:33–45.

34. Mehlisch D, Brown P. Single-dose therapy with diclofenac potassium, aspirin, or placebo following dental impaction surgery. *Today's Ther Trends* 1994;**12**:15–31.

35. Ahlström U, Bakshi R, Nilsson P, Wåhlander L. The analgesic efficacy of diclofenac dispersible and ibuprofen in postoperative pain after dental extraction. *Eur J Clin Pharmacol* 1993;**44**:587–8.

36. Bakshi R, Frenkel G, Dietlein G, Meurer Witt B, Schneider B, Sinterhauf U. A placebo-controlled comparative evaluation of diclofenac dispersible versus ibuprofen in postoperative pain after third molar surgery. *J Clin Pharmacol* 1994;**34**:225–30.

37. Bakshi R, Jacobs L, Lehnert S, Picha B, Reuther J. A double-blind, placebo-controlled trial comparing the analgesic efficacy of two formulations of diclofenac in postoperative dental pain. *Curr Ther Res Clin Exp* 1992;**52**:435–42.

38. Henry D, Lim LL, Rodriguez LAG, Gutthann SP, Carson JL, Griffin M, *et al.* Variability in risk of gastrointestinal complications with individual non-steroidal anti-inflammatory drugs: results of a collaborative meta-analysis. *BMJ* 1996;**312**:1563–6.

Studies included

Ahlström U, Bakshi R, Nilsson P, Wåhlander L, 1993. The analgesic efficacy of diclofenac dispersible and ibuprofen in postoperative pain after dental extraction. *Eur J Clin Pharmacol*;**44**:587–8.

Arnold JD, 1990. Ketoprofen, ibuprofen, and placebo in the relief of postoperative pain. *Adv Ther*;**7**:264–75.

Bakshi R, Jacobs LD, Lehnert S, Picha B, Reuther J, 1992. A double-blind, placebo-controlled trial comparing the analgesic efficacy of two formulations of diclofenac in postoperative dental pain. *Curr Ther Res*;**52**:435–42.

Bakshi R, Frenkel G, Dietlein G, Meurer Witt, B, Schneider B, Sinterhauf U, 1994. A placebo-controlled comparative evaluation of diclofenac dispersible versus ibuprofen in postoperative pain after third molar surgery. *J Clin Pharmacol*;**34**:225–30.

Cooper SA, Needle SE, Kruger GO, 1977. Comparative analgesic potency of aspirin and ibuprofen, *J Oral Surg*;**35**:898–903.

Cooper SA, Engel J, Ladov M, Precheur H, Rosenheck A, Rauch D, 1982. Analgesic efficacy of an ibuprofen codeine combination. *Pharmacotherapy*;**2**:162–7.

Cooper SA, Berrie R, Cohn P, 1988. Comparison of ketoprofen, ibuprofen, and placebo in a dental surgery pain model. *Adv Ther*;**5**:43–53.

Cooper SA, Schachtel BP, Goldman E, Gelb S, Cohn P, 1989. Ibuprofen and acetaminophen in the relief of acute pain: a randomized, double blind, placebo controlled study. *J Clin Pharmacol*;**29**:1026–30.

Forbes JA, Barkaszi BA, Ragland RN, Hankle JJ, 1984. Analgesic effect of fendosal, ibuprofen and aspirin in postoperative oral surgery pain. *Pharmacotherapy*;**4**:385–91.

Forbes JA, Kehm CJ, Grodin CD, Beaver WT, 1990. Evaluation of ketorolac, ibuprofen, acetaminophen, and an acetaminophen codeine combination in postoperative oral surgery pain. *Pharmacotherapy*;**10**:94–105S.

Forbes JA, Beaver WT, Jones KF, Kehm CJ, Smith WK, Gongloff CM, *et al.,* 1991. Effect of caffeine on ibuprofen analgesia in postoperative oral surgery pain. *Clin Pharmacol Ther*;**49**:674–84.

Forbes JA, Edquist IA, Smith FG, Schwartz MK, Beaver WT, 1991. Evaluation of bromfenac, aspirin, and ibuprofen in postoperative oral surgery pain. *Pharmacotherapy*;**11**:64–70.

Forbes JA, Beaver WT, Jones KF, Edquist IA, Gongloff CM, Smith WK, *et al.,* 1992 Analgesic efficacy of bromfenac, ibuprofen, and aspirin in postoperative oral surgery pain. *Clin Pharmacol Ther*;**51**:343–52.

Frame JW, Evans CR, Flaum GR, Langford R, Rout PG, 1989. A comparison of ibuprofen and dihydrocodeine in relieving pain following wisdom teeth removal. *Br Dent J*;**166**:121–4.

Fricke JR, Halladay SC, Francisco CA, 1993. Efficacy and safety of naproxen sodium and ibuprofen for pain relief after oral surgery. *Curr Ther Res*;**54**:619–27.

Gay C, Planas E, Donado M, Martinez JM, Artigas R, Torres F, *et al.,* 1996. Analgesic efficacy of low doses of dexketoprofen in the dental pain model: a randomised, double-blind, placebo-controlled study. *Clin Drug Invest*;**11**:320–30.

Hebertson RM, Storey N, 1994. The comparative efficacy of diclofenac potassium, aspirin, and placebo in the treatment of patients with pain following gynecologic surgery. *Today's Ther Trends*;**12**:33–45.

Heidrich G, Slavic Svircev V, Kaiko RF, 1985. Efficacy and quality of ibuprofen and acetaminophen plus codeine analgesia. *Pain*;**22**:385–97.

Hersh EV, Cooper S, Betts N, Wedell D, MacAfee K, Quinn P, *et al.,* 1993. Single dose and multidose analgesic study of ibuprofen and meclofenamate sodium after third molar surgery. *Oral Surg Oral Med Oral Pathol*;**76**:680–7.

Hersh EV, Ochs H, Quinn P, MacAfee K, Cooper SA, 1993. Narcotic receptor blockade and its effect on the analgesic response to placebo and ibuprofen after oral surgery. *Oral Surg Oral Med Oral Pathol*;**75**:539–46.

Jain AK, Ryan JR, McMahon FG, Kuebel JO, Walters PJ, Noveck C, 1986. Analgesic efficacy of low dose ibuprofen in dental extraction pain. *Pharmacotherapy*;**6**:318–22.

Jain AK, Mcmahon FG, Ryan JR, Narcisse C, 1988. A double-blind study of ibuprofen 200 mg in combination with caffeine 100 mg, ibuprofen 400 mg, and placebo in episiotomy pain. *Curr Ther Res*;**43**:762–79.

Kiersch TA, Halladay SC, Koschik M, 1993. A double-blind, randomized study of naproxen sodium, ibuprofen, and placebo in postoperative dental pain. *Clin Ther*;**15**:845–54.

Laska EM, Sunshine A, Marrero I, Olson N, Siegel C, McCormick N, 1986. The correlation between blood levels of ibuprofen and clinical analgesic response. *Clin Pharmacol Ther*;**40**:1–7.

Laveneziana D, Riva A, Bonazzi M, Cipolla M, Migliavacca S, 1996. Comparative efficacy of oral ibuprofen arginine and intramuscular ketorolac in patients with postoperative pain. *Clin Drug Invest*;**11**:8–14.

McQuay HJ, Angell K, Carroll D, Moore RA, Juniper RP, 1996. Ibuprofen compared with ibuprofen plus caffeine after third molar surgery. *Pain*;**66**:247–51.

Mehlisch DR, Brown P, 1994. Single-dose therapy with diclofenac potassium, aspirin, or placebo following dental impaction surgery. *Today's Ther Trends*;**12**:15–31.

Mehlisch DR, Sollecito WA, Helfrick JF, Leibold DG, Markowitz R, Schow CE, *et al.,* 1990. Multicenter clinical trial of ibuprofen and acetaminophen in the treatment of postoperative dental pain. *J Am Dent Assoc*;**121**:257–63.

Mehlisch DR, Jasper RD, Brown P, Korn SH, McCarroll K, Murakami AA, 1995. Comparative study of ibuprofen lysine and acetaminophen in patients with postoperative dental pain. *Clin Ther*;**17**:852–60.

Nelson S, Brahim J, 1994. An evaluation of the analgesic efficacy of diclofenac potassium, aspirin, and placebo in postoperative dental pain. *Today's Ther Trends*;**12**:3–14.

Nelson SL, Brahim JS, Korn SH, Greene SS, Suchower LJ, 1994. Comparison of single-dose ibuprofen lysine, acetylsalicylic acid, and placebo for moderate-to-severe postoperative dental pain. *Clin Ther*;**16**:458–65.

Pagnoni B, Ravanelli A, Degradi L, Rossi R, Tiengo M, 1996. Clinical efficacy of ibuprofen arginine in the management of postoperative pain associated with suction termination of pregnancy. A double-blind placebo-controlled study. *Clin Drug Invest*;**11**:27–32.

Parker DA, Gibbin KP, Noyelle RM, 1986. Syrup formulations for post tonsillectomy analgesia: a double blind study comparing ibuprofen, aspirin and placebo. *J Laryngol Otol*;**100**:1055–60.

Schachtel BP, Thoden WR, Baybutt RI, 1989. Ibuprofen and acetaminophen in the relief of postpartum episiotomy pain. *J Clin Pharmacol*;**29**:550–3.

Seymour RA, Hawkesford JE, Weldon M, Brewster D, 1991. An evaluation of different ibuprofen preparations in the control of postoperative pain after third molar surgery. *Br J Clin Pharmacol*;**31**:83–7.

Seymour RA, WardBooth P, Kelly PJ, 1996. Evaluation of different doses of soluble ibuprofen and ibuprofen tablets in postoperative dental pain. *Br J Oral Maxil Surg*;**34**:110–14.

Sunshine A, Olson NZ, Laska EM, Zighelboim I, de Castro A, de Sarrazin C, 1983. Ibuprofen, zomepirac, aspirin, and placebo in the relief of postepisiotomy pain. *Clin Pharmacol Ther*;**34**:254–8.

Sunshine A, Roure C, Olson N, Laska EM, Zorrilla C, Rivera J, 1987. Analgesic efficacy of two ibuprofen codeine combinations for the treatment of post-episiotomy and postoperative pain. *Clin Pharmacol Ther*;**42**:374–80.

Chapter 10

Comparison of the analgesic efficacy of NSAIDs given by different routes in acute and chronic pain

Summary

The aim of this review was to test the evidence for a difference in analgesic efficacy and adverse effects of NSAIDs administered by different routes. The relevant published RCTs were comparisons of the same drug given by different routes. Presence of internal sensitivity was sought as a validity criterion. Analgesic and adverse effect outcomes were summarised and synthesised qualitatively.

In 26 trials (2225 patients analysed) eight different NSAIDs were tested in 58 comparisons. In 15 trials (58%) different routes for the same drug were compared. Drugs were administered by intravenous, intramuscular, intrawound, rectal and oral routes in postoperative pain (14 trials), renal colic (4), acute musculoskeletal pain (1), dysmenorrhoea (1) and rheumatoid arthritis (6). Five of the 15 direct comparisons were invalid because they reported no difference between routes but without evidence of internal sensitivity.

In all three direct comparisons in renal colic, intravenously administered NSAID had a faster onset of action than intramuscular or rectal. In one direct comparison in dysmenorrhoea, orally administered NSAID was better than rectal. In the five direct comparisons in postoperative pain the results were inconsistent. In one direct comparison in rheumatoid arthritis intramuscular NSAID was better than oral. Injected and rectal administration had some specific adverse effects.

In renal colic there is evidence that NSAIDs act quickest when given intravenously. This may be clinically relevant. In all other pain conditions there is a lack of evidence of any difference between routes. In pain conditions other than renal colic, there is, therefore, a strong argument in favour of giving NSAIDs orally rather than by injection.

This chapter of the review has been published in full by Tramèr and colleagues.[1]

Introduction

Oral NSAIDs are an important component of simple 'low-technology' pharmacological control of both acute and chronic pain. Oral NSAIDs can be surprisingly effective in patients with moderate to severe postoperative pain. Compared with placebo, oral ibuprofen, 400 mg, will result in one in every three patients getting at least 50% relief of pain over 6 hours.[2] This is a high standard of effectiveness and one against which more complicated methods of delivering adequate analgesia have to be judged. Invasive procedures like continuous extradural opiate infusion, or PCA, carry recognised risks[3] and may not be available or appropriate for the majority of patients with acute or chronic pain.

While oral NSAIDs can be effective, the advent of rectal and injectable formulations has led to a fashion for using these routes. This is reflected in the US Agency for Health Care and Policy Research Acute Pain Guidelines, in which the options for postoperative pain include systemic administration of NSAIDs with no mention of the oral route.[4] There are clinical circumstances in which use of the oral route is not possible, such as patients who cannot swallow, who are unconscious, nauseated or who have an ileus. If NSAIDs are indicated for such patients, then rectal or injectable formulations are the only options. However, in the much commoner circumstance of preoperative premedication of conscious day-case patients who can swallow, or in other acute and chronic pain conditions, are there any reasons for using rectal or injected formulations rather than oral? The reasons could be greater efficacy, or similar efficacy with fewer adverse effects. However, the evidence for any such advantage over oral use is unclear and, in this review, the existing evidence from published reports of direct comparisons of NSAIDs given by different routes in both acute and chronic pain is assessed.

Methods

Full reports of published RCTs of direct comparisons of NSAIDs administered by different routes and tested in acute or chronic pain with pain outcomes were sought. A number of different search strategies[5] were used to identify eligible reports in MEDLINE (Knowledge Server®, Silver Platter, 1966–96), EMBASE (1986–96) and the Oxford Pain Relief Database (1950–93). The terms 'NSAID', 'non-steroidal anti-inflammatory', and individual drug names were used in conjunction with 'postoperative pain', 'renal colic', '*colic', 'intravenous', 'intramuscular', and 'rectal' in searching, including combinations and without restriction in language. Additional reports were identified from the reference lists of retrieved reports, review articles and textbooks, by hand-searching locally available anaesthesia journals, and by contacting pharmaceutical companies with licensed parenteral or rectal NSAID preparations.

Excluded reports

Abstracts, letters, review articles and use of topical formulations (skin, mucous membranes, eye) or intra-articular use were not considered. Unpublished reports were not sought. Reports in which the numbers of patients per treatment group were less than ten were excluded. Authors were not contacted.

Included reports

Each report which could possibly meet the inclusion criteria was read by at least two authors independently and scored for inclusion and methodological quality using a validated 3-item, 5-point scale.[6] Authors met to agree scores. Reports which were described as 'randomised' were given 1 point, and a further point if the method of randomisation was described and adequate (such as a table of random numbers). There was a *pre hoc* agreement that trials without concealment of treatment allocation (allocation according to patient's date of birth, for instance) would be excluded from further analysis because of the well-documented risk of overestimation of treatment effects in such trials.[7,8] One point was given when the trial was described as 'double-blind'. When the method of double-blinding was described and adequate (double-dummy method, for instance), a further point was given. Finally, reports which described the numbers of and reasons for withdrawals were given 1 point. Thus, the maximum score for an included RCT was 5 and the minimum score was 1.

Data extraction and analysis

These trials compared drug efficacy across different routes of administration. Therefore, the primary focus was on trials which compared the same drug given by different routes. Only such direct comparisons were regarded as relevant to this review. Comparisons of different drugs across routes were regarded as irrelevant and were not analysed. Comparisons between NSAID and non-NSAID controls (opioid, placebo) were not considered.

Each trial was checked for specific design details with potential impact on trial validity. These details were first, whether or not the design included internal sensitivity measures, either a negative control (placebo or no treatment) or at least two dose levels of an active drug. There was a prior agreement that trials which reported equivalence (i.e. no difference) between routes but which had no index of internal sensitivity would be regarded as invalid and not considered for data synthesis. Second, the extent to which blinding was protected by using a double-dummy design was checked. Finally, baseline pain intensity was recorded in trials in which pain was treated (chronic pain settings, for instance), and pain intensity without analgesic intervention was recorded in prophylaxis trials (such as postoperative setting).

Information on the clinical setting, inclusion criteria, number of patients, study design, and drugs, route and doses used was extracted from the reports, together with information on analgesic measurements and results, and adverse effects. Analgesic efficacy was estimated by extracting data of significant difference ($p < 0.05$, as reported in the original trials) between NSAID arms. Relevant outcomes were pain intensity at rest or on movement, and additional analgesic consumption. Quantitative analysis of combined data was proposed. There was a prior hypothesis that there was no clinically relevant difference between routes of administration with NSAIDs and, specifically, that the oral route would be no different from the other routes of administration.

Results

A total of 26 RCTs (2225 patients analysed), published between 1970 and 1996, were considered eligible for the review *(Table 32)*.

Of the 26 trials, 14 were in postoperative pain (1268 patients),[9–22] four in renal colic (647 patients),[23–26] one in acute musculoskeletal pain (77 patients),[27] one in dysmenorrhoea (32 patients),[28] and six in rheumatoid arthritis (201 patients).[29–34] Different doses of eight different NSAIDs (diclofenac, ibuprofen, indomethacin,

TABLE 32 *RCTs comparing analgesic efficacy of NSAIDs administered by different routes*

Study	Quality score (1–5)	Regimen: drug, dose, route (no. of patients)	Number of patients	Setting	Pain outcomes	Internal sensitivity		Double-dummy	Overall efficacy *(> better than; < worse than; = similar)*	Adverse effects
						Index	Given			
		[Bold: relevant trials, i.e. same drug across route]				**[Bold: fulfilled validity criteria in relevant trials]**				
Postoperative										
Parke, *et al.*, 1995[9]	5	**Ketorolac, 30 mg i.v., + saline i.m. (38)** **Ketorolac, 30 mg i.m., + saline i.v. (38)** Saline i.m. + i.v. (37)	113	Major orthopaedic (moderate or severe pain)	Time to onset of analgesia: NSD. Time to first subsequent analgesic: ketorolac i.v. = i.m. > placebo. Number achieving 1-point decrease of pain scale within 30 minutes: ketorolac i.m. = placebo; i.v. > placebo. Patients' rating: ketorolac i.v.> placebo.	**Placebo**	**Yes**	**Yes**	**i.v. = i.m.**	No serious adverse events.
Hyrkas, *et al.*, 1992[10]	2	**Diclofenac, 50 mg p.o. rapid + 100 mg p.o. retard premedication (20 mins) (51)** **Diclofenac, 50 mg i.m. + 100 mg p.o. retard pre-medication (20 mins) (51)** Placebo (49)	151	Third molar	PR 0–8 hours (VAS mean): diclofenac p.o. and i.m. significantly better than placebo. Rescue analgesics: significantly less needed in diclofenac groups.	**Placebo**	**Yes**	**Yes**	**p.o. + p.o. = p.o. + i.m.**	Diclofenac p.o. 17/51; diclofenac i.m. 16/51; placebo 16/49.
Ben-David, *et al.*, 1996[11]	2	**Ketorolac, 30 mg i.m. induction (14)** **Ketorolac, 30 mg i.w. induction (14)** **Ketorolac 30 mg i.v. induction (14)** **Ketorolac, 30 mg p.o. premedi-cation (1 hour) (14)** No treatment (14)	90	Inguinal hernia repair (local anaesthesia)	Dipyrone at 90 minutes: i.m. 2/14; i.w. 3/14; i.v. 1/14; p.o. 7/14 (significant); control 10/14. Buprenorphine at 90 minutes: i.m., i.w., i.v., and p.o. 0/14; control 5/14. VAS PI supine/sitting (90 minutes): i.m., i.w., and i.v. significantly better than p.o. and control.	**No treatment control**	**Yes**	No	**i.m. = i.v. = i.w. > p.o.**	Stated as none.
Campbell, *et al.*, 1990a[12]	1	**Diclofenac, 1 mg/kg i.v. induction (40)** **Diclofenac, 1 mg/kg i.m. induction (40)** Fentanyl, 1 µg/kg i.v. induction (40) Saline i.m. induction (40)	160	Third molar	Analgesic needs 'nil': diclofenac i.v. 14/40; diclofenac i.m. 12/40; fentanyl 3/40; saline 5/40. VAS PI (mean) at 30 minutes postoperatively (60 minutes postinjection): diclofenac i.v. significantly better than all other groups.	**Placebo**	**Yes**	No	**i.v. > i.m.**	Bleeding time intra-operational (30 minutes post-injection): significantly increased with i.m. diclofenac. No other adverse effects reported.
Jakobsson, *et al.*, 1996[13]	2	**Diclofenac, 75 mg i.m. premedication (10–20 mins) (50)** **Diclofenac, 50 mg p.o. premedication (10–20 mins) (50)** Ketorolac, 30 mg i.m. premedication (10–20 mins) (50) Saline, i.m. premedication (10–20 mins) (50)	200	Minor gynaecology	No pain (discharge): diclofenac i.m. 43/50 and p.o. 34/50; ketorolac i.m. 44/50; placebo 34/50 (i.m. NSAID vs. p.o. or placebo significant difference). No analgesics (discharge): diclofenac i.m. 37/50 and p.o. 27/50; ketorolac i.m. 38/50; placebo 27/50.	**Placebo**	**Yes**	No	**i.m. > p.o.**	Emesis: no difference. Anxiety: significantly less in i.m. groups. No other adverse effects reported.

continued

TABLE 32 *contd* *RCTs comparing analgesic efficacy of NSAIDs administered by different routes*

Study	Quality score (1–5)	Regimen: drug, dose, route (no. of patients)	Number of patients	Setting	Pain outcomes	Internal sensitivity		Double-dummy	Overall efficacy *(> better than; < worse than; = similar)*	Adverse effects
						Index	Given			
		[Bold: relevant trials, i.e. same drug across route]				**[Bold: fulfilled validity criteria in relevant trials]**				
Postoperative contd										
Moore, *et al.*, 1993[14]	1	**Diclofenac, 75 mg i.m. postoperative + 75 mg/12 hours: total 300 mg/ 48 hours (16) Diclofenac, 100 mg p.r. postoperative + 100 mg/12 hours: total 500 mg/ 48 hours (16)**	32	Thoracotomy	PI (VAS mean): i.m. vs. p.r., NSD. Analgesic consumption (mean mg papaveretum): NSD.	No	No	No	i.m. = p.r.	i.m. 1/16; p.r. 5/16 (2 diarrhoea).
Dennis, *et al.*, 1995[15]	3	Diclofenac, 100 mg p.r. premedication (1 hour) (20) Ketorolac, 10 mg i.v. induction (20)	40	Knee arthroscopy	Pain after 24 hours (nil or mild): ketorolac i.v. 11/20; diclofenac p.r. 14/20. Activity restriction (none or mild): ketorolac i.v. 12/20; diclofenac p.r. 16/20.	No	No	Yes	i.v. = p.r.	No adverse effects reported.
Lysak, *et al.*, 1994[16]	4	Ketorolac, 60 mg i.m. premedication (30 minutes) (29) Piroxicam, 40 mg p.o. premedication (90 minutes) (28) Fentanyl, 100 µg i.v. induction + 2 x 25 µg i.v. intra-operative (27)	84	Gynaecology laparoscopy	PI (mild at discharge): ketorolac i.m. 90%; piroxicam p.o. 97%; fentanyl 63%. Morphine required in postanaesthetic care unit: ketorolac i.m. 16/29; piroxicam p.o. 25/28; fentanyl 20/27.	No	No	Yes	i.m. = p.o.	Ketorolac i.m. 7/29; piroxicam p.o. 18/28; fentanyl 8/27.
Morrison, *et al.*, 1994[17]	3	Paracetamol, 650 mg p.o. postoperative (30 minutes) (20) Ibuprofen, 600 mg p.o. postoperative (30 minutes) (20) Ketorolac, 60 mg i.v. end of surgery (20)	60	Strabismus (adult)	VAS PI at 5 hours: ketorolac i.v. significantly better than paracetamol and ibuprofen p.o. No additional analgesia at 5 hours: paracetamol 4/20; ibuprofen p.o. 0/20; ketorolac i.v. 13/20.	No	Yes	No	i.v. > p.o.	Nausea: ketorolac i.v. 0/20, paracetamol 2/20, ibuprofen 2/20. No other adverse effects mentioned.
Morley-Forster, *et al.*, 1993[18]	3	Ketorolac, 30 mg i.m. induction (31) Indomethacin, 100 mg p.r. induction (31) Placebo i.m. and p.r. (25)	87	Gynaecology or breast	VAS PI: significantly lower with NSAID at 15 and 90 minutes but not at 60 minutes (p.r. = i.m.). No additional analgesics: both NSAIDs significantly better than placebo.	Placebo	Yes	Yes	p.r. = .i.m	Ketorolac i.m. 7/31; indomethacin p.r. 9/31; placebo 8/25.
Murrell, *et al.*, 1996[19]	2	Indomethacin, 100 mg p.r. + saline i.m. induction (38) Ketorolac, 30 mg i.m. + placebo p.r. induction (51) Placebo p.r. and i.m. induction (48)	137	Gynaecology laparoscopy	Analgesic use up to 180 minutes (fentanyl, paracetamol/codeine): NSD. VAS PI: no difference at 30 and 60 minutes. *Post hoc* significant difference between ketorolac and placebo at 180 minutes.	Placebo	No	Yes	p.r. = i.m.	No difference in frequency of complaints of pain at injection site.
Roelofse, *et al.*, 1993[20]	1	Tenoxicam, 20 mg i.v. intra-operative + 20 mg p.o. postoperative (12) Diclofenac, 75 mg i.m. intra-operative + 50 mg p.o. postoperative (13)	25	Third molar	VAS PI: significantly lower with diclofenac at 1, 2 and 3 hours.	No	Yes	No	(i.m. + p.o.) > (i.v. + p.o.)	Discomfort due to i.m. injection: 13/13 with diclofenac.

continued

TABLE 32 *contd* *RCTs comparing analgesic efficacy of NSAIDs administered by different routes*

Study	Quality score (1–5)	Regimen: drug, dose, route (no. of patients)	Number of patients	Setting	Pain outcomes	Internal sensitivity		Double-dummy	Overall efficacy *(> better than; < worse than; = similar)*	Adverse effects
						Index	Given			
		[Bold: relevant trials, i.e. same drug across route]				**[Bold: fulfilled validity criteria in relevant trials]**				
Postoperative contd										
Rusy, *et al.*, 1995[21]	2	Ketorolac, 1 mg/kg i.v. induction (25) Paracetamol, 35 mg/kg p.r. induction (25)	50	Tonsillectomy (children)	'Objective pain score' (blood pressure, crying, agitation, movement, verbal report): ketorolac > paracetamol at 2 hours. No difference at 30 minutes, 1 and 3 hours. Additional analgesics up to 3 hours: morphine and codeine, NSD; paracetamol: significantly less with ketorolac.	No	No	No	i.v. = p.r.	Extra homeostatic measurements: significantly more with ketorolac (not related to route).
Wakeling, *et al.*, 1996[22]	4	Diclofenac, 100 mg p.r. premedication (1 hour) + placebo p.o. (19) Piroxicam, 40 mg p.o. premedication (1 hour) + placebo p.r. (20)	39	Third molar	Median time to rescue analgesic: diclofenac 350 minutes; piroxicam 305 minutes ($p > 0.05$). No analgesic after 18 hours: diclofenac 2/19; piroxicam 6/20. VAS PI similar at any time.	No	No	Yes	p.o. = p.r.	Vomiting: piroxicam 3/20; diclofenac 0/19. Nausea: piroxicam 2/20; diclofenac 1/19.
Renal colic										
Muriel-Villoria, *et al.*, 1995[23]	4	**Dipyrone, 1 g i.m. + placebo i.v. (71)** **Dipyrone, 1 g i.v. + placebo i.m. (30)** **Dipyrone, 2 g i.v. + placebo i.m. (76)** **Dipyrone, 2 g i.m. + placebo i.v. (71)** **Diclofenac, 75 mg i.m. + placebo i.v. (32)** **Diclofenac, 75 mg i.v. + placebo i.m. (22)**	302	Renal colic (VAS PI > 50)	Proportion of patients with > 50% improvement: significant differences dipyrone, 2 g i.v. >1 g i.v. at 10 minutes; diclofenac, 75 mg i.v. > i.m. at 20 minutes; dipyrone, 2 g i.v. > 2 g i.m. at 10 and 20 minutes; dipyrone, 1 g i.v. > 1 g i.m. at 20 minutes.	**Dose response**	**Yes**	**Yes**	**i.v. > i.m.; i.v. > i.m.; i.v. > i.m.**	i.v.: vomiting x 1 (diclofenac). i.m.: drowsiness x 3 (1 dipyrone, 1 g; 1 dipyrone, 2 g; 1 diclofenac); drowsiness x 1 (diclofenac).
Nelson, *et al.*, 1988[24]	1	**Indomethacin, 100 mg p.r. (47)** **Indomethacin, 50 mg i.v. (37)**	84	Renal colic	VAS PI at 10 minutes: i.v. significantly lower than p.r.; at 30 minutes, NSD. Supplementary analgesics: p.r. 16/47; i.v. 8/37.	No	**Yes**	**Yes**	**i.v. > p.r.**	30/84 non-drug-related withdrawals (non-retention of suppositories and others). Drug-related: p.r. 8/47; i.v. 18/37.
Nissen, *et al.*, 1990[25]	3	**Indomethacin, 100 mg p.r. (63)** **Indomethacin, 50 mg i.v. (53)**	116	Renal colic	VAS PI at 10 and 20 minutes: i.v. significantly lower than p.r.; at 30 minutes, NSD. Supplementary analgesics: p.r. 17/63; i.v. 5/53 ($p = 0.03$).	No	**Yes**	No	**i.v. > p.r.**	i.v. 44/53; p.r. 29/63 ($p = 0.03$).
El-Sherif, *et al.*, 1990[26]	2	Indomethacin, 50 mg i.v. (44) Diclofenac, 50 mg i.m. (47) Avafortan (dipyrone + antispasmodic) i.v. (54)	145	Renal colic	PR ('complete') after first dose at 30 minutes: indomethacin i.v. 37/44; diclofenac i.m. 31/47; avafortan 45/54. NSAID i.v. significantly better than i.m.	No	Yes	No	i.v. > i.m.	Indomethacin i.v. 5/44; diclofenac i.m. 3/44; avafortan 0/54.

continued

TABLE 32 *contd* *RCTs comparing analgesic efficacy of NSAIDs administered by different routes*

Study	Quality score (1–5)	Regimen: drug, dose, route (no. of patients)	Number of patients	Setting	Pain outcomes	Internal sensitivity		Double-dummy	Overall efficacy (*> better than; < worse than; = similar*)	Adverse effects
						Index	Given			
		[Bold: relevant trials, i.e. same drug across route]				**[Bold: fulfilled validity criteria in relevant trials]**				
Acute musculoskeletal pain										
Turturro, *et al*, 1995[27]	5	Ketorolac, 60 mg i.m. + placebo p.o. (40) Ibuprofen, 800 mg p.o. + placebo i.m. (37)	77	Acute musculo-skeletal pain (treatment)	VAS PI 0–120 minutes: NSD between ketorolac i.m. and ibuprofen p.o.	No	No	Yes	i.m. = p.o.	Dyspepsia: ketorolac 1/40; ibuprofen 2/37. Sedation: ketorolac 1/40; ibuprofen 0/37. Dry mouth: ketorolac 0/40; ibuprofen 1/37.
Dysmenorrhoea										
Ylikorkala, *et al.*, 1980[28]	4	Crossover (n = 32) **Naproxen, 500 mg p.o. 6-hourly + placebo p.r. Naproxen, 500 mg p.r. 6-hourly + placebo p.o.**	32	Primary dysmenorrhoea	No difference in number of additional analgesics taken. Spasmodic PR (score): significantly better with p.o. All other symptoms (score): NSD. Patients' overall assessment: NSD.	No	**Yes**	**Yes**	**p.o. > p.r.**	p.r. irritation after naproxen suppositories: 2/32.
Rheumatoid arthritis										
Dougados, *et al.*, 1992[29]	3	**Ketoprofen, 2 x 50 mg p.o. + placebo i.m. (20) Ketoprofen, 100 mg i.m. + placebo p.o. (20)**	40	Rheumatoid arthritis (VAS PI > 40)	Decrease in VAS PI, patients' global judgement, maximum decrease in VAS PI after treatment: p.o. better than i.m. (NSD). Delay until lowest pain intensity: i.m. shorter than p.o. ($p < 0.05$).	No	**Yes**	**Yes**	**i.m. > p.o.**	Active p.o. (6/20): 3 nausea, 1 vomiting, 1 indigestion, 1 diarrhoea, 1 headache, 2 vertigo, 1 dyspnea. Active i.m. (5/20): 1 headache, 2 somnolence, 1 pruritus, 1 vertigo.
Baber, *et al.*, 1980[30]	4	Crossover (n = 13) Placebo p.r. + placebo p.o. for 1 week, then **Indomethacin, 100 mg p.r. + placebo p.o. for 1 week Indomethacin, 100 mg p.o. + placebo p.r. for 1 week**	13	Rheumatoid arthritis	Articular index, grip strength, pain score (analogue 0–9), morning stiffness, digital joint size, rescue analgesics (paracetamol) and patients' preference: NSD between p.o. and p.r.	No	No	Yes	p.o. = p.r.	p.o.: 17/13, pr: 11/13. Diarrhoea: 4/13 p.r. Indigestion: 5/13 p.o., 2/13 p.r.
Hansen, *et al.*, 1984[31]	3	Crossover (n = 12) Placebo p.r. + placebo p.o. for 2 weeks, then **Indomethacin, 100–150 mg p.o. + placebo p.r. for 2 weeks Indomethacin, 100–150 mg p.r. + placebo p.o. for 2 weeks**	12	Rheumatoid arthritis	Morning stiffness and pain (PI on 50 mm VAS): no difference between p.r. and p.o.	No	No	Yes	p.o. = p.r.	Endoscopy: equal amounts gastric mucosal damage. p.r. discomfort with suppositories: 8 p.o. active + p.r. placebo; 7 p.r. active + p.o. placebo; 7 p.o. and p.r. placebo.
Huskisson, *et al.*, 1970[32]	3	Crossover (n = 20) **Indomethacin, 100 mg p.r. + placebo p.o. for days 1 and 3 Indomethacin, 100 mg p.o. + placebo p.r. for days 2 and 4**	20	Rheumatoid arthritis	Patients' preference: p.o. significantly better than p.r. No difference in pain, morning stiffness, duration of stiffness.	No	No	Yes	p.o. = p.r.	Nausea, anorexia, epigastric discomfort: p.o. 2/20; p.r. 2/20.

continued

TABLE 32 *contd* *RCTs comparing analgesic efficacy of NSAIDs administered by different routes*

Study	Quality score (1–5)	Regimen: drug, dose, route (no. of patients)	Number of patients	Setting	Pain outcomes	Internal sensitivity		Double-dummy	Overall efficacy *(> better than; < worse than; = similar)*	Adverse effects
						Index	Given			
		[Bold: relevant trials, i.e. same drug across route]				**[Bold: fulfilled validity criteria in relevant trials]**				
Rheumatoid arthritis contd										
Iversen, *et al.*, 1981[33]	4	Crossover (n = 22) **Indomethacin SR, 75 mg p.o. + placebo p.r. for 2 weeks Indomethacin, 100 mg p.r. + placebo p.r. for 2 weeks**	22	Rheumatoid arthritis	Day and night pain (4-point scale), morning stiffness (minutes): NSD. Conventional grip strength: p.r. significantly better than p.o. (6 mm).	No	No	Yes	p.o. = p.r.	p.o. (5): headache, dizziness, 'loose bowel motions', stomach pain. p.r. (8): headache, 'loose bowel motions'.
Uddenfeldt, *et al.*, 1993[34]	4	Crossover (n = 94) Ketoprofen CR, 200 mg p.o. + placebo p.r. for 3 weeks Indomethacin, 100 mg p.r. + placebo p.o. for 3 weeks	94	Rheumatoid arthritis	Morning stiffness (duration): p.o. significantly better than p.r. Awakenings during night: p.r. significantly better than p.o. NSD in pain on awakening (VAS), articular index, patients' and doctors' assessment (7-step scale), rescue analgesics.	No	No	Yes	p.o. = p.r	p.r. irritation: indomethacin 7/94. Gastro-intestinal problems: keto-profen 40/94; indomethacin 28/94.

ketoprofen, ketorolac, naproxen, piroxicam, tenoxicam), given by intravenous, intramuscular, intrawound, rectal and oral routes, were tested in 58 single-dose or multiple-dose comparisons.

The median quality score of all trials was 3 (range 1–5). Quantitative analysis was not considered appropriate because of the variety of clinical settings, drugs, doses, routes and pain outcomes reported. Instead, any statistically significant difference between treatments was extracted from the original reports and documented in table format as had been done previously for other qualitative systematic reviews.[35,36] A 'vote counting' procedure was then agreed, giving positive or negative votes if there was evidence of presence or absence, respectively, of a significant difference between routes.

In all, 15 trials (58% of all analysed trials) were relevant to this review because the same drug was compared by different routes of administration.[9–14,23–25,28–33] In nine of them (35% of all trials), the same drug was compared at the same dose.[9,11,12,23,28–32]

In five of the 15 relevant trials equivalence was reported between routes but there was no index of internal sensitivity.[14,30–33] These trials were not, therefore, analysed further.

Of the ten relevant trials in which a significant difference between routes was reported, or equivalence was reported but with an index of internal sensitivity (i.e. which were valid), five were in postoperative pain,[9–13] three were in renal colic,[23–25] one was in dysmenorrhoea,[28] and one was in rheumatoid arthritis.[29] Six of them used a double-dummy design.[9,10,23,24,28,29]

Postoperative pain

Of 14 trials in postoperative pain, five were valid direct comparisons. They compared diclofenac or ketorolac across routes.

In one trial, diclofenac, 1 mg/kg injected intravenously at induction of anaesthesia, led to significantly lower pain intensity scores 30 minutes after surgery than the same dose given intramuscularly at induction.[12] In two other trials no difference was found between ketorolac, 30 mg, given either intravenously or intramuscularly at induction.[9,11] In one of these trials, inguinal hernia repair was performed under local anaesthesia with very low pain scores during the postoperative observation period whether or not an NSAID or no treatment was given.[11] In the same trial, both intramuscular and intravenous ketorolac, 30 mg, at induction led to significantly lower pain scores and less rescue analgesics at 90 minutes after surgery than the same dose taken orally but 1 hour before surgery.[11] Group sizes in this trial were small (i.e. 14 patients per group) and it was not of double-dummy design.

In another trial with larger groups (50 patients per group) but, again, not of double-dummy design, less pain and rescue analgesics at discharge were

reported with diclofenac, 75 mg, intramuscularly compared with the same drug given orally but at a lower dose (50 mg).[13] In yet another trial, diclofenac, 150 mg, taken orally was compared with diclofenac, 50 mg, intramuscularly plus 100 mg orally.[10] The drugs were given as premedication using a double-dummy design, and group sizes were large (50 patients per group). No difference was found between the two forms of administration.

Renal colic

Of four trials in renal colic, there were three valid direct comparisons.[23–25] Dipyrone, diclofenac, and indomethacin given by different routes were compared. Two were of double-dummy design.[23,24] In one, baseline pain before treatment was started (at least 50 mm on a 100 mm VAS) was defined.[23] Group sizes in these three trials were between 22 and 76 patients.

In one trial, pain relief was tested with dipyrone, 1 g or 2 g, and diclofenac, 75 mg, given intramuscularly compared with intravenously.[23] At 10 and 20 minutes after administration, the proportion of patients with at least 50% improvement was significantly in favour of the intravenous route with each drug and dose.

In the two other trials, intravenous indomethacin, 50 mg, was compared with the same drug given rectally but at double the dose.[24,25] Despite the intravenous dose being only half the rectal dose, significant improvement (less pain intensity, fewer rescue analgesics) was reported for the intravenous compared with the rectal route. Again these differences were apparent only at 10 or 20 minutes post-administration.

Dysmenorrhoea

Only one trial in dysmenorrhoea was found.[28] This crossover trial in 38 patients compared oral with rectal naproxen, both 500 mg at 6-hourly intervals, using a double-dummy design. Relief of spasmodic pain was significantly better with the oral route.

Chronic pain

Six trials in rheumatoid arthritis were found. Five of them were direct comparisons but only one small trial using a double-dummy design with 20 patients per group was valid.[29] Patients with defined baseline pain (at least 40 mm on a 100 mm scale) receiving ketoprofen, 100 mg, intramuscularly reported a significantly shorter delay until the lowest pain intensity score was achieved than patients receiving the same dose of the same drug orally.

Other pain conditions

No direct comparisons from other pain conditions were found.

Adverse effects

Commonly reported adverse effects independent of the route of administration were nausea, vomiting, dizziness, drowsiness, sedation, anxiety, dyspepsia, indigestion, and dry mouth *(Table 32)*. Two studies reported bleeding time changes.[12,21] In 12 patients with rheumatoid arthritis treated for 2 weeks with indomethacin, 100–150 mg orally and rectally, respectively, in a study of crossover design, endoscopically diagnosed gastric mucosal damage was independent of the route of administration.[31]

Adverse effects related to the route of administration were most often reported for intramuscular and rectal regimens *(Table 32)*. Discomfort at the site of injection[19,20] was the most frequent complaint relating to intramuscular injections. After rectal administration, diarrhoea,[14,30] rectal irritation,[28,31,34] and non-retention of suppositories[24] were reported.

Comment

Many doctors use injected or rectal NSAIDs when the oral route could be used. This is despite advice to use the least invasive route possible, with the statement that no study has specifically compared the analgesic efficacy of alternative routes of the same drug.[37] Reasons for preferring injected or rectal formulations when the oral route could be used might be greater efficacy or faster onset of pain relief. The safety argument would be that these efficacy benefits were achieved at no greater (or acceptably greater) level of adverse effects. Patients may prefer oral to rectal dosing.[38] There are also legal ramifications, because of the obligation for consent if drugs are given rectally while a patient is asleep.[39]

Using evidence from systematically searched published reports of RCTs with direct comparisons, we wanted to compare the benefits and risks of NSAIDs given by different routes in acute and chronic pain. Systematic reviews are powerful instruments to gain more insight in treatment efficacy and harm. Ideally, dichotomous outcomes would be extracted from original reports and combined using biostatistical methods. In some circumstances, dichotomous data may be extracted from measurements which were originally not binary outcomes.[40]

However, such quantitative analysis was not possible because of the variety of clinical settings, drugs, doses, routes and pain outcomes reported.

Systematic reviewers then have to rely on statistically significant results as reported in the original reports and apply a vote-counting procedure.[35] It is obvious that such a qualitative approach is vulnerable to bias. Vote counting take no account of the size of the trial or of the size of any difference in effect. In such analyses, pre-set validity criteria become especially important in order to minimise the risk of bias.[36]

These trials highlighted different methodological problems affecting the validity of the trials.

First, if the null hypothesis was that there was no difference between the routes, comparisons of the same drug given by different routes might be expected. It might be desirable to concentrate only on comparisons of the same drugs at the same dose. However, only about half of these trials compared the same NSAID across routes, thus addressing what was regarded as the clinically relevant question. Only one-third of all trials would have satisfied stricter validity rules (i.e. direct comparison of the same drug at the same dose).

Second, the classical approach to design of analgesic trials is to build in an index of internal sensitivity, either by using a placebo (or no treatment) control, or by including a high and a low dose of a standard analgesic to establish a dose–response relationship. What such designs seek to achieve is a defence against equivalence of treatments. Lack of internal sensitivity is a key issue in equivalence trials.[41] This has been shown in systematic reviews of analgesic trials.[36] Without such controls, equivalence in a comparison of two or more drugs may mean that the methods of measuring pain or its relief failed in that study, rather than that this was a true negative result of no difference between the analgesic effect of the drugs. Power calculations cannot be a defence against methodological failure. Only a positive result (significant difference) despite the lack of negative controls is an adequate vindication of such methods.

Eight (31%) of the 26 trials has a method built in that ensured internal sensitivity in the form of a placebo control,[9,10,12,13,18,19] a no-treatment control,[11] or two dose levels of the same drug given by the same route.[23] Five direct comparisons reported equivalence between routes but had no index of internal sensitivity. These trials were, therefore, invalid and were excluded from further analysis.

Third, although all these trials were, by definition, comparisons of NSAIDs given by different routes, only 17 (65%) were of double-dummy design. The blinding of the other trials must be questioned. In trials with deficient blinding, the therapeutic effect may be exaggerated.[7] While all trials in rheumatoid arthritis used a double-dummy technique, this was true for only half of the surgical trials. An extreme example was the comparison of an oral drug given 1 hour before surgery with the same drug given by injection at induction.[11] This trial did not use a double-dummy method and reported better analgesic efficacy for the parenteral route compared with the oral.

Finally, a *pre hoc* defined pain intensity, sufficient to provide measurable change after study treatment, was reported in only a minority of trials measuring pain relief.[9,23,29] Very low pain intensity scores independent of the treatment were reported in some trials where pain was meant to be prevented, such as in a surgical setting. If there is no pain, analgesia cannot be measured. A pain trial without an adequate baseline pain intensity is not a valid assay.[36]

Applying our rules of validity to these 26 trials revealed that only 15 (58%) of all systematically searched trials were relevant to this review (i.e. investigated the same drug given by different routes). Five of them had to be excluded because their results could not be interpreted (i.e. they reported equivalence but had no index of internal sensitivity). This meant that only ten trials (38% of all trials) could be analysed, and only six of them were of double-dummy design. In renal colic, there was evidence from three valid direct comparisons that the intravenous route acted significantly faster than the rectal or intramuscular route. Although this difference was only evident during the first 10 to 20 minutes, the faster onset of action is likely to be clinically relevant in this specific setting. In the only trial in dysmenorrhoea, one outcome measurement indicated that oral NSAID may be better than rectal. In one trial in rheumatoid arthritis, one isolated endpoint suggested that intramuscular ketoprofen may be superior to oral; however, the clinical utility of this is unclear. Finally, in the surgical setting, results were far from being conclusive.

Reporting of adverse effects was generally poor and mostly not related to route of administration. Rectal and intramuscular routes were most likely to have specific local adverse effects. These have to be taken into account when the advantages of one route over the other are discussed.

With the exception of the renal colic setting, these trials constitute a lack of evidence for any

difference rather than evidence of lack of difference between NSAIDs given by different routes. This is not just semantics; if there is adequate evidence of a lack of difference then practice should change, reverting to the safest and simplest option, the oral route. If there is a lack of evidence (rather than evidence of a lack of difference), then a research agenda is set, to determine whether or not there is any clinical advantage of one route over another. It could be argued that patients should again receive the safest and simplest option unless they agree to participate in a randomised comparison of different routes of administration. The research agenda should be to design simple comparisons of the same drug at the same dose across route, with validity and, ideally, with standardised outcome measures in the various studies to make combined quantitative analysis possible.

It does not seem right that over 2200 patients have already participated in trials over the past 26 years and, yet, for the majority of clinical settings we still cannot answer the simple question, 'Is it better to give NSAIDs by injection or suppository than to take them orally?'

References

1. Tramèr M, Williams J, Carroll D, Wiffen PJ, McQuay HJ, Moore RA. Systematic review of direct comparisons of non-steroidal anti-inflammatory drugs given by different routes for acute pain. *Acta Anaesth Scand* 1997;**42**:71–9.

2. Collins SL, Moore A, McQuay HJ, Wiffen PJ. Oral ibuprofen and diclofenac in postoperative pain: a quantitative systematic review. *Eur J Pain*. In press.

3. Bates DW, Cullen DJ, Laird N, Petersen LA, Small SD, Servi D, *et al.* Incidence of adverse drug events and potential adverse drug events. *JAMA* 1995;**274**:29–34.

4. Acute Pain Management Guideline Panel. Acute Pain Management: operative or medical procedures and trauma. Clinical Practice Guideline No. 1. Agency for Health Care Policy and Research, US Department of Health and Human Services. AHCPR Publication No 92-0032. Rockville, MD: Public Health Service, 1992:24–5.

5. Jadad AR, McQuay HJ. A high-yield strategy to identify randomized controlled trials for systematic reviews. *Online J Curr Clin Trials* [serial online] 1993;Doc No 33:3973 words;39 paragraphs;5 tables.

6. Jadad AR, Moore RA, Carroll D, Jenkinson C, Reynolds DJM, Gavaghan DJ, *et al.* Assessing the quality of reports of randomized clinical trials: is blinding necessary? *Control Clin Trials* 1996;**17**:1–12.

7. Schulz KF, Chalmers I, Hayes RJ, Altman DG. Empirical evidence of bias: dimensions of methodological quality associated with estimates of treatment effects in controlled trials. *JAMA* 1995;**273**:408–12.

8. Carroll D, Tramèr M, McQuay H, Nye B, Moore A. Randomization is important in studies with pain outcomes: systematic review of transcutaneous electrical nerve stimulation in acute postoperative pain. *Br J Anaesth* 1996;**77**:798–803.

9. Parke TJ, Millett S, Old S, Goodwin APL, Rice ASC. Ketorolac for early postoperative analgesia. *J Clin Anesth* 1995;**7**:465–9.

10. Hyrkas T, Ylipaavalniemi P, Oikarinen VJ, Hampf G. Postoperative pain prevention by a single-dose formulation of diclofenac producing a steady plasma concentration. *J Oral Maxillofac Surg* 1992;**50**:124–7.

11. Ben-David B, Baune-Goldstein U, Goldik Z, Gaitini L. Is preoperative ketorolac a useful adjunct to regional anesthesia for inguinal herniorrhaphy? *Acta Anaesthesiol Scand* 1996;**40**:358–63.

12. Campbell WI, Kendrick R, Patterson C. Intravenous diclofenac sodium. Does its administration before operation suppress postoperative pain? *Anaesthesia* 1990;**45**:763–6.

13. Jakobsson J, Rane K, Davidson S. Intramuscular NSAIDs reduce post-operative pain after minor outpatient anaesthesia. *Eur J Anaesth* 1996;**13**:67–71.

14. Moore AP, Thorpe JA, Mulley BA, Cruise M. Which diclofenac form is best? Post-thoracotamy pain: intramuscular v. rectal administration of diclofenac. *Hosp Pharmacy Pract* 1993;September:431–2,435.

15. Dennis AR, Leeson-Payne CG, Hobbs GJ. A comparison of diclofenac with ketorolac for pain relief after knee arthroscopy. *Anaesthesia* 1995;**50**:904–6.

16. Lysak SZ, Anderson PT, Carithers RA, DeVane GG, Smith ML, Bates GW. Postoperative effects of fentanyl, ketorolac, and piroxicam as analgesics for outpatient laparoscopic procedures. *Obstet Gynecol* 1994;**83**:270–5.

17. Morrison NA, Repka MX. Ketorolac versus acetaminophen or ibuprofen in controlling postoperative pain in patients with strabismus. *Ophthalmology* 1994;**101**:915–18.

18. Morley-Forster P, Newton PT, Cook M. Ketorolac and indomethacin are equally efficacious for the relief of minor postoperative pain. *Can J Anaesth* 1993;**40**:1126–30.

19. Murrell GC, Leake T, Hughes PJ. A comparison of the efficacy of ketorolac and indomethacin for postoperative analgesia following laparoscopic surgery in day patients. *Anaesth Intens Care* 1996;**24**:237–40.

20. Roelofse JA, van der Bijl P, Joubert JJD. An open comparative study of the analgesic effects of tenoxicam and diclofenac sodium after third molar surgery. *Anesth Pain Control Dentistry* 1993;**2**:217–22.

21. Rusy LM, Houck CS, Sullivan LJ, Ohlms LA, Jones DT, McGill TJ, *et al.* A double-blind evaluation of ketorolac tromethamine versus acetaminophen in pediatric tonsillectomy: analgesia and bleeding. *Anesth Analg* 1995;**80**:226–9.

22. Wakeling HG, Barry PC, Butler PJ. Postoperative analgesia in dental day case surgery. A comparison between Feldene 'Melt' (piroxicam) and diclofenac suppositories. *Anaesthesia* 1996;**51**:784–6.

23. Muriel-Villoria X, Muriel-Villoria C, Zungri-Telo E, *et al.* Comparison of the onset and duration of the analgesic effect of dipyrone, 1 or 2 g, by the intramuscular or intravenous route, in acute renal colic. *Eur J Clin Pharmacol* 1995;**48**:103–7.

24. Nelson CE, Nylander C, Olsson AM, Olsson R, Pettersson BA, Wallstrom I. Rectal v. intravenous administration of indomethacin in the treatment of renal colic. *Acta Chir Scand* 1988;**154**:253–5.

25. Nissen I, Birke H, Olsen J, Wurtz E, Lorentzen K, Salomon H, *et al.* Treatment of ureteric colic. Intravenous versus rectal administration of indomethacin. *Br J Urol* 1990;**65**:576–9.

26. El-Sherif A, Foda R, Norlen L, Yahia H. Treatment of renal colic by prostaglandin synthetase inhibitors and avafortan (analgesic antispasmodic). *Br J Urol* 1990;**66**:602–5.

27. Turturro MA, Paris PM, Seaberg DC. Intramuscular ketorolac versus oral ibuprofen in acute musculoskeletal pain. *Ann Emerg Med* 1995;**26**:117–20.

28. Ylikorkala O, Puolakka J, Kauppila A. Comparison between naproxen tablets and suppositories in primary dysmenorrhea. *Prostaglandins* 1980;**20**:463–8.

29. Dougados M, Listrat V, Duchesne L, Amor B. Efficacite comparee du ketoprofene en fonction de sa voie d'administration (intra-musculaire ou per os). Une etude en double aveugle contre placebo au cours de la polyarthrite rhumatoide [Comparative efficacy of ketoprofen related to the route of administration (intramuscular or per os). A double-blind study versus placebo in rheumatoid arthritis]. *Rev Rhum Mal Osteoartic* 1992;**59**:769–73.

30. Baber N, Sibeon R, Laws E, Halliday L, Orme M, Littler T. Indomethacin in rheumatoid arthritis: comparison of oral and rectal dosing. *Br J Clin Pharmacol* 1980;**10**:387–92.

31. Hansen TM, Matzen P, Madsen P. Endoscopic evaluation of the effect of indomethacin capsules and suppositories on the gastric mucosa in rheumatic patients. *J Rheumatol* 1984;**11**:484–7.

32. Huskisson EC, Taylor RT, Burston D, Chuter PJ, Hart FD. Evening indomethacin in the treatment of rheumatoid arthritis. *Ann Rheum Dis* 1970;**29**:393–6.

33. Iversen O, Fowles M, Vlieg M, Smidt N, Wigley R. Sustained release indomethacin: a double blind comparison with indomethacin suppositories. *N Z Med J* 1981;**93**:261–2.

34. Uddenfeldt P, Leden I, Rubin B. A double-blind comparison of oral ketoprofen 'controlled release' and indomethacin suppository in the treatment of rheumatoid arthritis with special regard to morning stiffness and pain on awakening. *Curr Med Res Opin* 1993;**13**:127–32.

35. Carroll D, Tramèr M, McQuay H, Nye B, Moore A. Transcutaneous electrical nerve stimulation in labour pain: a systematic review. *Br J Obstet Gynaecol* 1997;**104**:169–75.

36. Kalso E, Tramèr M, Carroll D, McQuay H, Moore RA. Pain relief from intra-articular morphine after knee surgery: a qualitative systematic review. *Pain* 1997;**71**:642–51.

37. Cashman J, McAnulty G. Nonsteroidal anti-inflammatory drugs in perisurgical pain management. Mechanisms of action and rationale for optimum use. *Drugs* 1995;**49**:51–70.

38. Vyvyan HAL, Hanafiah Z. Patients' attitudes to rectal drug administration. *Anaesthesia* 1995;**50**:983–4.

39. Mitchell J. A fundamental problem of consent. *BMJ* 1995;**310**:43–8.

40. McQuay HJ, Moore RA. Using numerical results from systematic reviews in clinical practice. *Ann Intern Med* 1997;**126**:712–20.

41. Jones B, Jarvis P, Lewis JA, Ebbutt AF. Trials to assess equivalence: the importance of rigorous methods. *BMJ* 1996;**313**:36–9.

Chapter 11

Topically-applied NSAIDs

Summary

In this chapter the effectiveness and safety of topical NSAIDs in acute (soft tissue trauma, strains and sprains) and chronic pain conditions (osteoarthritis, tendinitis) are assessed. In all, 86 RCTs involving 10,160 patients were found. Measures approximating at least 50% pain relief, local and systemic adverse effects were extracted. Analysis was undertaken at 1 week for acute and 2 weeks for chronic conditions using relative benefit and NNT.

In acute pain conditions, placebo-controlled trials had a relative benefit of 1.7 (95% CI, 1.5–1.9) and an NNT of 3.9 (95% CI, 3.4–4.4). Analysing by drug (at least three trials), ketoprofen (NNT 2.6), felbinac (3.0), ibuprofen (3.5) and piroxicam (4.2) had significant efficacy. Benzydamine and indomethacin were not distinguished from placebo.

In chronic pain conditions, placebo-controlled trials had a relative benefit of 2.0 (95% CI, 1.5–2.7) and an NNT of 3.1 (95% CI, 2.7–3.8). Small trials (< 40 treated patients) exaggerated the effectiveness of topical NSAIDs in acute conditions only (by 24%). There was no relationship between trial quality and treatment effect.

In both acute and chronic pain, local and systemic adverse events and drug-related study withdrawal had a low incidence and were no different from placebo. Topical NSAIDs are effective in relieving pain in acute and chronic conditions.

Introduction

Some topical NSAIDs are available without prescription and are widely advertised for acute and chronic painful conditions. In the UK, some 20–24 million (predominantly oral) NSAID prescriptions are written each year, 5% of total NHS prescriptions, with many more available without prescription. The attributable risk of going to hospital with gastrointestinal problems is between 1.3% and 1.6% annually for regular users of oral NSAIDs.[1] This raises the question whether using oral NSAIDs is worse than the disease for some patients.[2] Despite licensed status there is scepticism that topical NSAIDs have any action other than as rubefacients.[2,3] This systematic review was undertaken to examine the evidence that topical NSAIDs are effective and safe, and to determine whether there is evidence of differences between topical preparations.

Methods

Reports were sought of RCTs of topical noNSAIDs in which pain was an outcome. Reports were included which compared topical NSAID(s) with placebo, with another topical NSAID, or with an oral NSAID. A number of different search strategies in MEDLINE (1966–September 1996), EMBASE (1981–September 1996) and the Oxford Pain Relief Database (1950–94)[4] were used to locate reports, using individual drug name (generic and proprietary), together with the words 'administration', 'topical', 'gel', 'ointment', 'aerosol', 'cream', and combinations of these, without restriction on language. Additional reports were identified from the reference lists of retrieved reports and review articles. Librarians and medical directors of 12 pharmaceutical companies in the UK identified as marketing topical NSAIDs were asked for reports of RCTs of their products, including any unpublished reports. Abstracts were not sought. Authors were not contacted.

RCTs of NSAIDs with pain as an outcome in acute conditions (strains, sprains, sports injuries) or chronic conditions (arthritis, rheumatism) were included. Those in vaginitis, oral or buccal conditions, thrombophlebitis or experimental pain settings were not.

Reports were screened by two members of the team to eliminate those without pain outcomes, which were definitely not randomised, or were abstracts or reviews. Each report which could possibly be described as an RCT was read independently by each of the authors and scored using a 3-item, 1–5 score, quality scale.[5] Consensus was then achieved. The maximum score for an included RCT was 5 and the minimum 1.

Information about treatment(s) and control(s), condition studied, number of patients randomised and analysed, study design, observation periods,

outcome measures used for pain or global evaluation, analgesic outcome results, local skin irritation, systemic adverse effects and study withdrawal because of adverse effects was taken from each report by authors meeting to concur.

A clinically relevant outcome was defined as at least 50% pain relief. Only information that was available in dichotomous form was used for analysis. A hierarchy of measures was used for extraction which approximated, in this order of preference:

(i) patient global judgement (excellent/good)
(ii) pain on movement (no pain/slight pain)
(iii) spontaneous pain or pain at rest (no pain/slight pain)
(iv) physician global judgement (excellent/good) if defined against a stated scale.

The denominator was taken as the number of patients randomised, that is an intention-to-treat analysis. For acute conditions the effectiveness measure nearest to 1 week after start of treatment was taken and, for chronic conditions, 2 weeks. Prior hypotheses were that topical NSAIDs were no better than placebo and that there were no differences between them.

The scatter of success rates with topical NSAIDs against success rate with placebo[6] was used as a graphical means of exploring the consistency of efficacy and the homogeneity of the data. On such plots a scatter lying predominantly between the line of equality and the axis of the active intervention (topical non-steroidal) would suggest consistent efficacy with the intervention and relative homogeneity.

Relative risk or benefit with 95% CI was calculated for pain data using a random effects model[7] because the results were heterogeneous. Heterogeneity was assumed when $p > 0.1$. This was performed by pooling all data, by pooling data for an individual drug where there were at least three trials and, for sensitivity analysis, by quality score and treatment group size. A fixed effect model[8] was used for the (homogeneous) adverse effect data. A statistically significant improvement over control was assumed when the lower limit of the 95% CI of the relative benefit was > 1. NNTs and 95% CIs were calculated for effect data.[9,10] The NNT indicated how many patients with acute or chronic pain have to be treated with topical NSAIDs for one of them to achieve at least 50% pain relief who would not have done with placebo. A significant difference between NNTs was assumed when CIs did not overlap.

Results

The literature searches found 86 reports (10,160 patients) which fulfilled the inclusion criteria, 76 of which had dichotomous pain outcomes, including three unpublished reports with 1695 patients from a pharmaceutical company. The number of reports, patients, and the distribution of quality scores divided by acute or chronic, both placebo and active drug controlled, is shown in *Table 33*. Over 75% of placebo-controlled trials had quality scores of 3 or more. Conversely, 60% of active drug controlled trials had scores of 2 or less. Full details of trial design, outcome measures, and results are presented in *Tables 34–37*.

TABLE 33 *Number of reports, patients and the distribution of quality scores*

Trials	Number	Number of patients	Quality score (1–5)				
			1	2	3	4	5
Acute pain placebo-controlled	37	3556	1	6	10	13	7
Acute pain active drug controlled	24	4171	4	11	4	5	0
Chronic pain placebo-controlled	13	1161	0	3	5	5	0
Chronic pain active drug controlled	12	1272	2	5	3	2	0

Acute conditions

In all, 37 reports of 40 placebo-controlled trials of topical NSAIDs were found (see *Tables 34* and *35*). The mean size of group treated with topical drug was 47 patients (median 32). Studies were conducted in recent soft tissue injury, sprains, strains or trauma. Dichotomous pain outcomes were available for 1747 patients on active drug treatment and 1492 on placebo. An additional 24 reports of 24 trials compared different topical NSAIDs or formulations or route of administration in 4171 patients. In three studies, a topical NSAID was compared with oral; one such study also had a placebo control.

Relative benefit and 95% CIs are shown for each placebo controlled trial in *Figure 33*. Of the 37 comparisons, 27 showed statistical superiority for topical NSAIDs over placebo. The scatter of the proportion of patients with at least 50% pain relief with topical NSAID or placebo is shown in *Figure 34*. Of the 37 comparisons, 36 were in the

TABLE 34 *Placebo-controlled trials in acute painful conditions: trial design, outcome measures and results*

Study	Drug(s)	Condition	Numbers, study design and follow-up	Dosing regimen	Outcome measures	Analgesic outcome results	Skin irritation	Drug-related withdrawals and adverse effects	Quality score
Airaksinen, *et al.*, 1993	Ketoprofen, 2.5% gel; placebo gel	Acute soft tissue injuries, < 1 week	n = 56 parallel group 0, 3, 7 days	5 g twice daily	1. VAS PI on rest and movement 2. Patient and investigator global rating	1. Overall significant reduction in pain at rest with ketoprofen (NSD for placebo). 2. Significant difference ($p < 0.05$) in number of patients improved (patient global): 24/29 ketoprofen; 14/27 placebo.	5/29 ketoprofen; 4/27 placebo.	Withdrawals: 0/29 ketoprofen; 0/27 placebo. Adverse effects: 1/29 ketoprofen; 0/29 placebo.	2
Åkermark, *et al.*, 1990	Indomethacin, 1% spray; indomethacin, oral; placebo, spray/oral	Repetitive sports injuries	n = 70 parallel group double-dummy 3, 7, 14 days	Indomethacin spray, 0.5–1.5 ml three to five times daily; indomethacin 3 x 25 mg tablets	1. Improvement VAS 2. Physician global 5-point scale 3. Pain on movement, palpation, activity, 4-point scale	1. Indomethacin spray showed significant improvement ($p < 0.05$) days 3 and 7. 2. Marked improvement or symptom free at 1 week: 10/22 indomethacin spray; 5/23 indomethacin oral; 3/24 placebo. 3. Marked improvement or symptom free at 2 weeks: 16/22 indomethacin spray.	4/23 indomethacin spray; 0/23 indomethacin oral; 0/24 placebo.	Withdrawals: 1/23 indomethacin spray; 1/23 indomethacin oral; 0/24 placebo. Adverse effects: 4/23 indomethacin spray; 10/23 indomethacin oral; 0/24 placebo.	5
Aoki, *et al.*, 1984	Piroxicam, 0.5% gel; indomethacin, 1% gel; placebo, gel	Acute orthopaedic trauma	n = 252 multicentre, parallel group 0, 3, 7 days	1 g 3–4 times daily	Multiple outcomes: 1. Overall improvement patient 2. PI – movement, spontaneous	1. Significant difference in overall improvement ($p < 0.05$), piroxicam best. 2. Improvement day 7: piroxicam significantly better than placebo ($p < 0.01$). 3. Patients better or much better: 56/84 piroxicam; 41/84 indomethacin; 33/84 placebo.	1/84 piroxicam; 2/84 indomethacin; 2/84 placebo.	Withdrawals and adverse effects: 0/84 piroxicam; 0/84 indomethacin; 0/84 placebo.	4
Auclair, *et al.*, 1989	Niflumic acid, 2.5% gel; placebo, gel	Achilles heel tendinitis of recent origin	n = 243 parallel group 7, 21 days	5 g gel three times daily	1. Pain (VAS) on palpation 2. PI at dorsifexion 3. Global patient	1. Significantly more pain reduction than placebo. 2. Pain on dorsiflexion disappeared or improved: 75/117 niflumic acid; 69/110 placebo. 3. Global very good/good : 69/117 niflumic acid; 54/109 placebo.	5/123 niflumic acid; 6/116 placebo.	Adverse effect withdrawal: 1/123 niflumic acid; 0/116 placebo.	3
Baracchi, *et al.*, 1982	Ibuprofen, 10% cream; placebo, cream	Acute soft tissue trauma	n = 40 parallel group 3, 5, 7, 10, 12, 14 days	Twice daily	1. Categorical spontaneous pain, pain on movement and pressure 2. Investigator global	1. Ibuprofen significantly better than placebo (for spontaneous pain, $p < 0.001$). 2. Global (good or excellent response): 17/20 ibuprofen; 3/20 placebo.	No report of local effects.	Well tolerated.	4
Campbell & Dunn, 1994	Ibuprofen, 5% cream; placebo, cream	Acute ankle sprain, < 24 hours	n = 100 parallel group 2-week diaries	4 inches four times daily	VAS on rest and movement	1. Ibuprofen better than placebo on days 2 and 3. 2. Improved walking ability at day 7: 21/50 ibuprofen; 19/50 placebo.	Not reported.	Withdrawals: 0/50 ibuprofen; 0/50 placebo. Adverse effects: 1/50 ibuprofen; 0/50 placebo. 55/100 returned diaries.	4
Candela, *et al.*, 1986	Ketoprofen, gel; placebo, gel	Traumatic sport injuries	n = 30 parallel group 5, 10, 15 days	Twice daily	CAT scales, pain on pressure, on movement, functional limitation	1. Ketoprofen better than placebo. 2. Better/much better at day 10: 10/15 ketoprofen; 2/15 placebo.	Not reported.	Not reported.	1
Chaterjee, 1977	Benzydamine, 3% cream; placebo, cream	Soft tissue injuries	n = 51 parallel group 6 days	Three times daily	VRS pain – spontaneous, pressure, movement	1. Benzydamine better than placebo at 6 days spontaneous pain, pressure, movement. 2. None or slight pain on movement on day 6: 21/25 benzydamine; 12/25 placebo.	Not reported.	Withdrawals and adverse effects: 0/25 benzydamine; 0/25 placebo.	4

continued

TABLE 34 contd *Placebo-controlled trials in acute painful conditions: trial design, outcome measures and results*

Study	Drug(s)	Condition	Numbers, study design and follow-up	Dosing regimen	Outcome measures	Analgesic outcome results	Skin irritation	Drug-related withdrawals and adverse effects	Quality score
Diebshlag, 1986	Diclofenac, gel; placebo, gel	Ankle sprains	n = 20 crossover 2 x 1 week	*Ad libitum*	1. Ankle joint volume measurement 2. VAS PI	Reduced swelling and less pain with diclofenac.	Not reported.	Withdrawals and adverse effects: 0/20 diclofenac; 0/20 placebo.	3
Diebshlag, *et al.*, 1990	Ketorolac, 2% gel; etofenamate, 5% gel; placebo, gel	Acute ankle sprain	n = 37 parallel group 2, 3, 4, 8, 14 days	3 g three times daily	1. Ankle joint volume measurement 2. VAS PI	1. Ketorolac better than placebo and etofenemate. 2. Improved by day 3: 0/12 placebo; 11/13 ketorolac; 6/12 etofenamate.	1/13 ketorolac; 1/12 etofenamate; 0/12 placebo.	Withdrawals and adverse effects: 0/12 placebo; 0/13 ketorolac; 0/12 etofenamate.	4
Diebshlag & Knocker, 1987	Salicyclic acid, ointment; placebo, ointment	Ankle sprain, < 48 hours	n = 80 parallel group 2, 3, 4, 8, 15 days	10–15 cm ointment twice daily	1. Ankle joint volume measurement 2. VAS pain on rest and movement	Salicylic acid better than placebo for all measures.	Not reported.	Withdrawals and adverse effects: 0/40 salicylic acid; 0/40 placebo.	5
Dreiser, 1988	Ibuprofen, 5% cream; placebo, cream	Acute tendinitis, < 1 month	n = 64 parallel group 7 days	4 cm cream three times daily	VAS pain on rest, pressure, movement	1. Ibuprofen better than placebo ($p < 0.01$). 2. Global improvement: 26/32 ibuprofen; 13/28 placebo.	0/32 ibuprofen; 0/32 placebo.	Withdrawals and adverse effects: 0/32 ibuprofen; 0/32 placebo.	3
Dreiser, 1989	Ketoprofen, 5% gel; placebo, gel	Simple sprains	n = 60 parallel group 7 days	5 cm twice daily	VAS PI on rest, movement, patient global	Global improvement: 18/30 ketoprofen; 5/30 placebo.	0/30 ketoprofen; 1/30 placebo.	Withdrawals and adverse effects: 0/30 ketoprofen; 0/30 placebo.	5
Dreiser, *et al.*, 1990	Niflumic acid, 2.5% gel; placebo, gel	Uncomplicated ankle sprains	n = 60 parallel group 7 days	5 g gel three times daily	VAS PI, investigator PI, patient and investigator global	Patient global (improved or healed): 23/30 niflumic acid; 10/30 placebo.	3/30 niflumic acid; 1/30 placebo.	Withdrawals and adverse effects: 0/30 niflumic acid; 0/30 placebo.	5
Dreiser, *et al.*, 1994	Furbiprofen, patch 40 mg; placebo, patch	Ankle joint pain after post-traumatic strain	n = 131 parallel group 7 days	Two patches daily	VAS spontaneous pain by patient	Mean VAS significantly lower with furbiprofen. Pain better than moderate at day 7: 53/64 furbiprofen; 52/66 placebo.	Not reported.	Withdrawals: 0/64 furbiprofen; 0/66 placebo.	4
Fantato & de Gregorio, 1971	Benzydamine, 3% cream; placebo, cream	Oedema and post-traumatic pain	n = 52 parallel group 6 days	Three times daily	1. 4-point verbal rating 2. Investigator global with CAT scale	1. Benzydamine better than placebo. 2. ≥ 50% fall in symptom score: 22/26 benzydamine; 14/26 placebo.	Not reported.	Not reported.	5
Frahm, 1993	Salicylic acid, cream 2%; placebo, cream	Acute knee or ankle sprains	n = 156 parallel group 9 days	10 cm cream twice daily	VAS PI on movement and at rest	VAS PI significantly less on day 9 for salicylic acid cream.	0/78 salicylic acid; 0/78 placebo.	Withdrawals: 0/78 salicylic acid; 0/78 placebo. Adverse effects: 0/78 salicylic acid; 1/78 placebo.	5
Fujimaki, *et al.*, 1985	Piroxicam, 0.5% gel; indomethacin, 1% gel; placebo, gel	Musculoskeletal pain	n = 271 multicentre, parallel group 7 days	1 g three times daily	1. 4-point verbal rating scale pain on rest and movement 2. patient global	1. Both active treatments better than placebo in producing marked improvement. 2. Overall improvement better or much better: 44/92 piroxicam; 44/90 indomethacin; 40/89 placebo.	1/92 piroxicam; 6/90 indomethacin; 2/89 placebo.	Withdrawals: 0/92 piroxicam; 1/90 indomethacin; 0/89 placebo. Adverse effects: 0/92 piroxicam; 2/90 indomethacin; 0/89 placebo.	3
Haig, 1986	Benzydamine, 3% cream; placebo, cream	Acute soft tissue injuries	n = 43 parallel group 2, 4, 6 days	Six times daily	4-point verbal rating scale – spontanous pain and pain on movement	1. Pain on movement improved on day 6: NSD – 18/21 benzydamine; 13/22 placebo. 2. Spontanous pain improved on day 6: significantly different – 20/21 benzydamine; 14/22 placebo.	0/21 benzydamine; 0/22 placebo.	Withdrawals and adverse effects: 0/21 benzydamine; 0/22 placebo.	4
Julien, 1989	Ketoprofen, 0.5% gel; placebo, gel	Tendinitis	n = 60 parallel group 7 days	5 cm twice daily	1. Patient VAS PI on rest and movement 2. Overall patient assessment	Overall patient assessment (recovery, improvement) on day 7: 25/30 ketoprofen; 13/30 placebo.	1/30 ketoprofen; 0/30 placebo.	Withdrawals and adverse effects: 0/30 ketoprofen; 0/30 placebo.	4

continued

TABLE 34 contd *Placebo-controlled trials in acute painful conditions: trial design, outcome measures and results*

Study	Drug(s)	Condition	Numbers, study design and follow-up	Dosing regimen	Outcome measures	Analgesic outcome results	Skin irritation	Drug-related withdrawals and adverse effects	Quality score
Kockelbergh, *et al.*, 1985	Ketoprofen, 2.5% gel; placebo, gel	Acute soft tissue injuries, < 1 week	n = 74 parallel group baseline and 1 week	7.5 g gel twice daily	1. 4-point verbal rating VAS PI 2. Global	1. Ketoprofen better than placebo in producing improved symptoms. 2. Global, patient, good: 30/38 ketoprofen; 22/36 placebo.	1/38 ketoprofen; 1/36 placebo.	Withdrawals and adverse effects: 0/38 ketoprofen; 0/36 placebo.	2
Kockelbergh, *et al.*, 1985	Ketoprofen, 2.5% gel; placebo, gel	Acute low back pain, < 10 days	n = 40 parallel group 2 weeks	15 g gel with physiotherapy and ultrasound; 10 sessions	1. VAS pain and 5-point verbal rating 2. 3-point patient global rating	1. Significantly more patients with moderate/ severe pain at end in placebo group. 2. Global rating good: 13/20 ketoprofen; 9/20 placebo.	4/20 ketoprofen; 1/20 placebo.	Withdrawals and adverse effects: 1/20 ketoprofen; 0/20 placebo.	3
Lester, 1983	Salicylic acid, 2% cream; placebo, cream	Sprained ankle	n = 42 parallel group 7 days	Not stated	1. Ankle movement 2. Swelling 3. Pain 4. Return to normal activity	Pain relieved by 7 days: 18/20 salicylic acid; 13/22 placebo.	0/20 salicylic acid; 2/22 placebo.	Not reported.	3
Linde, *et al.*, 1985	Benzydamine, 5% cream; placebo, cream	Sprained ankle	n = 100 parallel group 8 days	Three times daily	1. Swelling 2. Pain on walking 3. Fit for work	Significant reduction in swelling with benzydamine; NSD for pain. Free of walking pain on day 8: 35/50 benzydamine; 40/50 placebo.	Not reported.	Not reported.	2
McLatchie, *et al.*, 1989	Felbinac, 3% gel; placebo, gel	Acute soft tissue injury	n = 231 parallel group baseline and 7 days	3 cm gel three times daily	1. VAS PI rest, movement, night pain 2. Investigator global	Good/very good treatment response (physician assessment): 85/118 felbinac; 46/113 placebo.	3/118 felbinac; 2/113 placebo.	Withdrawals and adverse effects: 0/118 felbinac; 0/113 placebo.	4
Morris, *et al.*, 1991	Felbinac, 3% gel; placebo, gel	Acute soft tissue sports injuries	n = 100 multicentre, parallel group baseline and 7 days	1 cm gel three times daily	1. Multiple global rating 2. VAS pain	1. Felbinac better than placebo. 2. Patients with good/very good results (patient global): 23/50 felbinac; 13/50 placebo.	0/50 felbinac; 0/50 placebo.	Withdrawals and adverse effects: 0/50 felbinac; 0/50 placebo.	4
Noret, *et al.*, 1987	Ketoprofen, 2.5% gel; placebo, gel	Minor sports injuries	n = 98 multicentre, parallel group 1, 3, 8 days	7.5 g gel twice daily	1. VAS PI 2. 4-point pain on pressure 3. Global	1. Ketoprofen better than placebo on many indices. 2. Global patient good or better: 39/51 ketoprofen; 9/47 placebo.	1/51 ketoprofen; 0/47 placebo.	Withdrawals and adverse effects: 1/51 ketoprofen; 0/47 placebo.	3
Parrini, *et al.*, 1992	Ketoprofen, 15% foam; placebo, foam	Soft tissue injuries	n = 169 parallel group 7 days	2 g three times daily	1. Categorical scale for spontaneous pain, on movement 2. Categorical scale global physician	1. Ketoprofen better than placebo for pain on pressure, movement and at rest. 2. Global physician excellent/good: 67/83 ketoprofen; 38/86 placebo.	0/83 ketoprofen; 0/86 placebo.	Withdrawals and adverse effects: 0/83 ketoprofen; 0/86 placebo.	4
Ramesh, *et al.*, 1983	Ibuprofen, 5% cream; placebo, cream	Soft tissue trauma	n = 80 parallel group 0, 3, 7, 10 days	5–10 cm three to four times daily	1. Pain on rest, pressure and movement; 4-point scale 2. Investigator global 3-point scale	Pain on movement on day 7 none/slight: 28/40 ibuprofen; 16/40 placebo.	1/40 ibuprofen; 1/40 placebo.	Withdrawals and adverse effects: 0/40 ibuprofen; 0/40 placebo.	4
Russell, 1991	Piroxicam, 0.5% gel; placebo, gel	Soft tissue injuries	n = 214 parallel group 7 days up to 21 days	1 g four times daily	1. VAS pain on rest and movement 2. Global 4-point scale 3. Daily pain charts	1. Piroxicam better than placebo at reducing pain by day 8. 2. Better joint mobility with piroxicam. 3. Global assessment good/ excellent: 79/100 piroxicam; 45/100 placebo.	4/102 piroxicam; 10/102 placebo.	Withdrawals: 1/102 piroxicam; 8/102 placebo. Adverse effects: 4/102 piroxicam; 7/102 placebo.	5

continued

TABLE 34 contd *Placebo-controlled trials in acute painful conditions: trial design, outcome measures and results*

Study	Drug(s)	Condition	Numbers, study design and follow-up	Dosing regimen	Outcome measures	Analgesic outcome results	Skin irritation	Drug-related withdrawals and adverse effects	Quality score
Sanguinetti, 1989	Biphenyl acetic acid, 3% gel; placebo, gel	Soft tissue traumas	n = 82 parallel group 7 days	Three times daily	Various scales	Global patient good/very good: 34/42 biphenyl acetic acid; 11/40 placebo.	0/42 biphenyl acetic acid; 0/40 placebo.	Adverse effects: 0/42 biphenyl acetic acid; 0/40 placebo.	4
Sinneger & Blanchard, 1981	Fentiazac, 5% cream; placebo, cream	Sport microtrauma	n = 20 parallel group 10 days	Twice or three times daily	Pain at rest, pressure, movement by physician	TOTPAR achieved within 10 days: 7/10 fentiazac; 1/10 placebo,	0/10 fentiazac; 0/10 placebo.	Withdrawals and adverse effects: 0/10 fentiazac; 0/10 placebo.	2
Taboada, 1992	Piroxicam, gel; placebo, gel	Acute musculo-skeletal pain	n = 40 parallel group 5–10 applications	Dose of drug and duration not stated; gels used with ultra-sound and infra-red treatment	Patient global	Excellent or good: 16/20 piroxicam; 6/20 placebo.	Not reported.	Not reported.	2
Thorling, *et al.*, 1990	Naproxen, 10% gel; placebo, gel	Sport injuries	n = 120 parallel group 7 days	2–6 times daily	1. Physician scoring of pain at rest, movement, swelling 2. Patient and physician global	Global patient, good/very good on day 7: 38/60 naproxen; 27/60 placebo.	1/60 naproxen; 0/60 placebo.	Withdrawals and adverse effects: 0/60 naproxen; 0/60 placebo.	3
Vecchiet & Colozzi, 1989	Meclofenamic acid, 5% gel; placebo, gel	Soft tissue injuries	n = 60 parallel group 5, 10 days	4 g twice daily	1. Categorical scale for spontaneous pain, pain on movement 2. Patient and physician global	1. Meclofenamic acid better than placebo. 2. Global patient: 30/30 meclofenamic acid; 19/30 placebo.	0/30 meclofenamic acid; 0/30 placebo.	Withdrawals and adverse effects: 0/30 meclofenamic acid; 0/30 placebo.	3
Wanet, *et al.*, 1979	Diethylamine salicylate; placebo, gel	Traumatic rheumatological injuries	n = 56 parallel group 15 days	Three times daily	Pain on rest and movement	Global assessment at end of treatment, good/very good: 20/32 salicylate; 9/24 placebo.	Not reported.	Not reported.	3
Zerbi, *et al.*, 1992	Ketoprofen, foam; ketoprofen, gel; placebo, foam	Painful traumatic injuries	n = 154 parallel group 7 days	Twice daily application, equivalent to 200 mg each time	1. Pain at rest, under pressure, movement 2. Global evaluation	1. Both active formulations significantly better than placebo. 2. Global patient (positive result): 33/46 foam; 35/49 gel; 13/42 placebo.	2/46 ketoprofen foam; 0/49 ketoprofen gel; 0/42 placebo.	Withdrawals: 0/46 ketoprofen foam; 0/49 ketoprofen gel; 1/42 placebo. No adverse effects withdrawals.	2

segment favouring treatment over placebo. The three trials which did not have dichotomous outcomes also reported statistical benefit for topical NSAID over placebo.

Pooled relative benefit for all 37 comparisons was 1.7 (95% CI, 1.5–1.9) and the NNT was 3.9 (95% CI, 3.4–4.3) *(Table 38)*. Pooling data only from trials with a quality score of at least 3 produced the same results. Sensitivity analysis by treatment group size showed that trials with a group size of at least 40 treated patients produced higher (worse) estimates for NNT of 4.8 (95% CI, 4.0–5.7) than all trials together. Trials with fewer than 40 treated patients produced a significantly lower (better) NNT of 2.6 (95% CI, 2.3–3.1) than either larger trials or all trials.

Pooling data for each drug studied in three or more trials showed ketoprofen, felbinac, ibuprofen and piroxicam to be statistically superior to placebo with NNTs ranging from 2.6 to 4.2. Indomethacin and benzydamine were no better than placebo *(Table 38)*.

The percentage of patients achieving at least 50% pain relief with active treatment or placebo in all studies in all trials (placebo and active drug controlled) in acute conditions is shown in *Figure 35* (lower panel). The range with placebo was from 0% to 80%. With topical NSAID it was from 30% to 100%. There was no significant difference in the (low) frequency of local or systemic adverse effects, or drug-related withdrawal *(Table 38)*.

Chronic conditions

The 13 placebo-controlled trials (see *Tables 36* and *37)* were predominantly in single joint arthritis and rheumatological disorders, with dichotomous outcomes from 547 patients on active drug treatment and 550 on placebo in 12 trials. In 12 other trials different NSAIDs were compared in 1272 patients. In two of these trials, topical NSAIDs were compared with oral.

TABLE 35 *Active drug controlled trials in acute painful conditions: trial design, outcome measures and results*

Study	Drug(s)	Condition	Numbers, study design and follow-up	Dosing regimen	Outcome measures	Analgesic outcome results	Skin irritation	Withdrawals and adverse effects	Quality score
Arioli, *et al.*, 1990	Piroxicam, 1% cream; diclofenac, 1% gel	Acute musculo-skeletal disorders	n = 75 parallel group, open design 3, 7, 14 days	1 g piroxicam cream, 4 g diclofenac gel, four times daily	1. Categorical and VAS scales for pain on movement, at rest, etc. 2. Patient global	1. Piroxicam better than diclofenac for some measures. 2. Patient global 7 days (better/much better): 34/38 piroxicam; 27/37 diclofenac.	0/38 piroxicam; 0/37 diclofenac.	0/38 piroxicam; 0/37 diclofenac.	2
Baixauli, *et al.*, 1990	Naproxen, 10% gel; ketoprofen, 10% gel	Acute soft tissue trauma < 24 hours	n = 30 parallel group 3, 7 days	5 cm naproxen, 3–5 cm ketoprofen, twice daily	1. Patient and investigator global rating 5-point 2. Improved 3-point	1. Cured or improved day 3: 10/15 naproxen; 12/14 ketoprofen. 2. Cured or improved day 7: 15/15 naproxen; 13/15 ketoprofen. 3. Patient global (good/ very good): 13/15 naproxen; 9/15 ketoprofen.	0/15 naproxen; 0/15 ketoprofen.	Withdrawals and adverse effects: 0/15 naproxen; 0/15 ketoprofen.	3
Bouchier-Hayes, *et al.*, 1990	Diclofenac, 1% gel; felbinac, 3% gel	Acute soft tissue injuries	n = 386 multicentre, parallel group 3, 7 days	4 g gel three times daily	VAS pain on rest pressure and movement	1. Diclofenac better than felbinac for some measures. 2. ≥ 50% improvement in pain on movement on day 7: 110/191 diclofenac; 100/195 felbinac.	Not reported.	0/191 diclofenac; 2/195 felbinac.	1 (probably not double-blind)
Butron, *et al.*, 1994	Naproxen, 10% gel; diclofenac, 1% gel	Sprains and contusions	n = 64 parallel group 4 days	As required	1. VAS pain on rest, movement 2. Patient and physician global	Patient global good/ excellent: 30/34 naproxen; 28/30 diclofenac.	3/34 naproxen; 4/30 diclofenac.	0/34 naproxen; 0/30 diclofenac.	2
Commandre, *et al.*, 1993	Niflumic acid, 2.5% gel; piroxicam, 0.5% gel	Acute sprains or tendinitis	n = 100 parallel group 7, 14 days	15 gm of each daily	1. Patient VAS 2. Investigator categorical 3. Patient global	1. Niflumic acid significantly better than piroxicam days 8 and 15. 2. Patient global day 8: 41/51 niflumic acid; 23/49 piroxicam.	3/51 niflumic acid; 4/49 piroxicam.	0/51 niflumic acid; 0/49 piroxicam. Adverse effects withdrawals: 1/51 niflumic acid; 1/49 piroxicam.	2
Curioni, *et al.*, 1985	Ibuproxam; ketoprofen; etofenamate	Soft tissue injuries	n = 60 parallel group 10 days	Twice daily application	1. Pain – spontaneous, on palpation, movement 2. Patient global	Some differences between groups.	2/20 ibuproxam; 3/20 ketoprofen; 1/20 etofenamate.	No information.	4
Diebschlag, *et al.*, 1992	Indomethacin, 1% gel (A); indomethacin, 1 % gel (B) (different vehicles)	Acute ankle sprain	n = 42 parallel group 2 weeks	Three times daily	Swelling, pain	No difference in swelling or pain between two preparations. Patient global, excellent or good: 19/19 (A); 21/22 (B).	0/19 (A); 1/22 (B).	No differences.	3
Gallachi, *et al.*, 1990	Diclofenac, 1% gel; diclofenac, 1.16% gel	Painful inflammatory symptoms	n = 50 parallel group 7, 14 days	2 g four times daily	1. Spontaneous pain 2. Pain on pressure 3. Patient global	1. NSD. 2. NSD. 3. 19/25 both groups good/excellent.	0/25 both groups.	No data.	2
Governali & Casalini, 1995	Ketoprofen, 5% gel; ketoprofen, 1% cream	Soft tissue injuries	n = 30 parallel group 7, 14 days	2–3 g of gel or cream three times daily	1. Pain – spontaneous, movement, pressure 2. Patient global	1. Gel significantly better than cream. 2. On day 7, excellent/ good: 14/15 ketoprofen gel; 9/15 ketoprofen cream.	0/15 ketoprofen gel; 0/15 ketoprofen cream.	Withdrawals and adverse effects: 0/15 ketoprofen gel; 0/15 ketoprofen cream.	2
Gualdi, *et al.*, 1987	Flunoxaprophene, gel; ketoprofen, gel	Soft tissue injuries	n = 60 parallel group 1, 4, 7, 10 days	3–5 cm of gel, twice daily	1. PI 2. Function 3. Patient global	NSD between groups.	1/30 flunox-aprophene; 3/30 ketoprofen.	No information.	2
Hallmeier & Michelbach, 1986	Etofenamate, 10% gel plus dressing; heparin/ dexpanthenol/ dimethyl-sulphoxide	Sports injuries	n = 60 parallel group 4 days	Not given	1. Oedema 2. Erythema 3. Movement 4. Patient global 'success'	4. Patient global: 26/30 etofenamate; 10/30 heparin.	0/30 etofenamate; 2/30 heparin.	Withdrawals: 0/30 etofenamate; 0/30 heparin.	2
Hallmeier, 1988	Etofenamate, 10% gel; diclofenac, 1% gel	Sprains and contusions	n = 60 parallel group, single blind 7 days	2–4 times daily	Patient global	Patient global very good/ good: 27/30 etofenamate; 13/30 diclofenac.	0/30 etofenamate; 0/30 diclofenac.	Withdrawals and adverse effects: 0/30 etofenamate; 0/30 diclofenac.	1

continued

TABLE 35 contd *Active drug controlled trials in acute painful conditions: trial design, outcome measures and results*

Study	Drug(s)	Condition	Numbers, study design and follow-up	Dosing regimen	Outcome measures	Analgesic outcome results	Skin irritation	Withdrawals and adverse effects	Quality score
Hosie, 1993	Felbinac, 3% foam; ibuprofen, 400 mg tablets	Acute lower back injury	n = 287 multicentre, parallel group, double-dummy 7, 14 days	2 g gel three times daily; 1 tablet three times daily	1. Pain 5-point scale 2. Investigator global	1. No difference between groups in symptom severity. 2. Both showed significant improvement. 3. No/mild pain on movement at 14 days: 99/140 felbinac foam; 109/147 ibuprofen oral.	1/140 felbinac foam; 3/147 ibuprofen oral.	25/140 felbinac foam; 19/147 ibuprofen oral.	4
Kroll, *et al.*, 1989	Piroxicam, 0.5% gel; diclofenac, 1.16% gel	Sprains and tendinitis	n = 173 parallel group, open to 14 days	1 g piroxicam, 2–4 g diclofenac, four times daily	1. Patient score of pain on movement (21-point VAS) 2. Patient global	Patient global excellent/ good: 63/84 piroxicam; 62/89 diclofenac.	4/84 piroxicam; 3/89 diclofenac.	Withdrawal: 2/84 piroxicam; 1/89 diclofenac. Adverse effects: 2/84 piroxicam; 0/89 diclofenac.	2
Montagna, *et al.*, 1990	Meclofenamic acid, 5% gel; naproxen, 10% gel	Painful musculo-skeletal disorders	n = 40 parallel group 4, 8, 15 days	Prescribed amounts twice daily	1. Pain – spontaneous and on movement 2. Patient global	1. No statistical difference between groups. 2. Excellent/good on day 8: 13/20 meclofenamic acid; 10/20 naproxen.	No data.	No data.	1
Oakland, 1993	Felbinac + placebo ultrasound; placebo gel + ultasound; felbinac + ultrasound	Acute injuries of lateral ankle ligaments	n = 220 parallel group days 3, 5, 7	1–2 g gel two to three times daily	1. Pain at rest 2. Investigator global	NSD.	3/147 felbinac; 3/73 placebo.	0/147 felbinac; 2/73 placebo.	3
Picchio, *et al.*, 1981	Ibuprofen, 10% gel; ketoprofen, 1% gel	Acute sports injuries	n = 40 parallel group 4, 8, 12, 16 days	Three times daily	5-point pain for pain at rest, on movement, spontaneous	1. Ibuprofen significantly better and faster than ketoprofen. 2. No pain on movement at 12 days: 16/20 ibuprofen; 10/20 ketoprofen.	Not reported.	0/20 ibuprofen; 0/20 ketoprofen.	3
Pineda, *et al.*, 1983	Felbinac, 3% gel; piroxicam gel, (?0.5%)	Acute soft tissue injuries	n = 172 multicentre, parallel group 3, 7 days	Felbinac, 180 mg/day, piroxicam, 18 mg/day, three times daily	1. Multiple, 10-point pain on rest, movement and night pain 2. Global 5-point	1. Complete recovery at 7 days: felbinac better than piroxicam ($p = 0.008$). 2. Good/very good global: 68/86 felbinac; 65/86 piroxicam.	5/86 felbinac; 1/86 piroxicam.	1/86 felbinac; 0/86 piroxicam.	1
Rosemeyer, 1991	Diclofenac, 1% gel; piroxicam, 0.5% gel	Distortion of ankle joint	n = 91 parallel group 3, 7, 10, 14 days	10 cm diclofenac, 3 cm piroxicam, four times daily	1. Pain at rest 2. Pain on pressure 3. Patient global	1 and 2. NSD at any time 3. Patient global excellent/ good: 35/44 diclofenac; 40/47 piroxicam.	4/44 diclofenac; 5/47 piroxicam.	Adverse effects withdrawal: 0/44 diclofenac; 1/47 piroxicam.	4
RPR1	Ketoprofen, gel; piroxicam, gel; diclofenac, gel	Acute soft tissue injury	n = 1575 parallel group 5 days	Ketoprofen, 4–5 g, piroxicam 1 g, diclofenac, 2–4 g, three times daily for 5 days	Patients' global assessment of injury	Greatly improved: 396/1048 ketoprofen; 69/263 piroxicam; 80/264 diclofenac.	Not reported.	Not reported.	2
Selligra & Inglis, 1990	Naproxen, 10% gel; flufenamic acid, 3% gel	Aoft tissue injuries	n = 100 parallel group, single-blind 7 days	2–6 times daily	Patients' global	Good/very good: 31/49 naproxen; 28/51 flufenamic acid.	1/49 naproxen; 0/51 flufenamic acid.	Withdrawals: 1/49 naproxen; 0/51 flufenamic acid. Adverse effects: 0/49 naproxen; 0/51 flufenamic acid.	2
Sugioka, *et al.*, 1984	Piroxicam, 0.5% gel; indomethacin, 1% gel	Non-traumatic disease of muscle or tendon	n = 366 multicentre, parallel group 1, 2 weeks	1 g three or four times daily	Multiple pain 4-point. 7-point symptom improvement	Piroxicam better than indomethacin. Patient self-assessment better/much better: 85/183 piroxicam; 55/183 indomethacin.	1/183 piroxicam; 12/183 indo-methacin.	Withdrawals: 1/183 piroxicam; 12/183 indomethacin. Adverse effects: 6/183 piroxicam; 26/183 indomethacin.	4
Tonutti, 1994	Ketoprofen, 5% gel; etofenamate, 5% gel	Soft tissue trauma	n = 30 parallel group 7 days	2–3 grams gel three times daily for up to 3 weeks	1. Pain, spontaneous, on movement, pressure 2. Patient global	1. Comparable efficacy. 2. Day 7 good/excellent: 10/15 ketoprofen; 11/15 etofenamate.	0/15 ketoprofen; 0/15 etofenamate.	0/15 ketoprofen; 0/15 etofenamate.	4
Vander-straeten & Scheumans, 1990	Etofenemate, 10% gel; naproxen, 275 mg tablets	Strains and sprains of lower limbs within 3 days	n = 60 parallel group 7, 17 days	5 cm gel, 1 tablet, three times daily	Categorical scales for spontaneous pain and pain on palpation	1. Day 7 no/slight pain: 13/30 etofenemate gel; 15/30 naproxen oral. 2. Clinical global good/ excellent improvement: 12/30 etofenemate gel; 13/30 naproxen oral.	1/30 etofenemate gel; 0/30 naproxen oral.	0/30 etofenemate gel; 6/30 naproxen oral. Withdrawals: 1/30 etofenemate gel; 2/30 naproxen oral.	2

TABLE 36 *Placebo controlled trials in chronic painful conditions: trial design, outcome measures and results*

Study	Drug(s)	Condition	Numbers, study design and follow-up	Dosing regimen	Outcome measures	Analgesic outcome results	Skin irritation	Withdrawals and adverse effects	Quality score
Algozzini, *et al.*, 1982	Trolamine salicylate, 10% cream; placebo, cream	Osteoarthritis of knee	n = 26 crossover 1 week	3.5 g cream, four times daily	1. 4-point PI 2. Numerical rating 0–10	1. NSD. 2. Patient preference: 8/26 salicylate; 6/26 placebo; 11/26 no preference. 3. PR from diaries: 9/26 salicylate; 6/26 placebo.	0/26 salicylate; 0/26 placebo.	Withdrawals and adverse effects: 0/26 salicylate; 0/26 placebo,	4
Bolten, 1991	Felbinac, 3% gel; placebo, gel	Acute extra-articular rheumatic disorders	n = 281 parallel group 0, 7, 14 days	1 g, 3 times daily	1. Categorical and VAS on rest and movement 2. VAS 3. Investigator global rating	1. Felbinac significantly better than placebo. 2. Global estimation of good/very good responses ($p < 0.001$): 67/142 felbinac; 39/139 placebo.	2/142 felbinac; 4/139 placebo.	Withdrawals and adverse effects: 0/142 felbinac; 0/139 placebo.	3
Camus, 1975	Diethylamine salicylate, cream; placebo, cream	Rheumatic disorders	n = 20 parallel group 10 days	Three times daily	4-point verbal rating	1. Salicylate better than placebo in giving relief over 10 days. 2. Pain reduced: 8/10 salicylate; 3/10 placebo.	0/10 salicylate; 0/10 placebo.	Withdrawals and adverse effects: 0/10 salicylate; 0/10 placebo.	2
Dreiser & Tisne-Camus, 1993	Diclofenac, plasters; placebo, plasters	Osteoarthritis of knee	n = 155 parallel group 4, 7, 15 days	Applied twice daily (each plaster contained 180 mg diclofenac derivative)	1. VAS 2. Global rating 5-point	1. Diclofenac better than placebo from day 4 2. Global rating excellent/good: 55/78 diclofenac; 21/77 placebo.	1/78 diclofenac; 3/77 placebo.	Withdrawals: 0/78 diclofenac; 0/77 placebo. Adverse effects: 0/78 diclofenac; 1/77 placebo.	4
El-Hadidi & El-Garf, 1991	Diclofenac, gel; ultrasound coupling gel	Painful rheumatic conditions	n = 120 parallel group 4 weeks	Three times per week	Physician judgement plus VAS PI by patient at rest and on movement Patient global	Diclofenac significantly better than regular coupling gel on all measures. Complete PR on passive movement at 2 weeks: 26/60 diclofenac; 18/60 regular.	2/60 diclofenac; 1/60 regular.	Withdrawals: 1/60 diclofenac; 0/60 regular. Adverse effects: 0/60 diclofenac; 0/60 regular.	3
Fotiades & Bach, 1976	Flufenamate, 3%, plus salicylate, 2%, gel; placebo, gel	Cervical, lumbar and shoulder pain and gonarthroses	n = 100 parallel group up to 20 days	Three or four times daily for 6–20 days	Point-scoring system including pain at rest, on pressure, pain relief, muscle spasm and movement	Scoring very good/good: 43/48 active drug; 26/52 placebo.	0/48 active drug; 0/52 placebo.	Withdrawals and adverse effects: 0/48 active drug; 0/52 placebo.	3
Galiazzi & Marcolongo, 1993	Diclofenac, plaster (slow-release); placebo, plaster	Rheumatological disorders	n = 60 parallel group 3, 5, 7, 14 days	Applied twice daily (each plaster contained 180 mg diclofenac derivative)	1. Multiple 4-point verbal rating and VAS 2. Investigator global scale	1. Diclofenac better than placebo in reducing pain. 2. Assessment of good/excellent response: 26/30 diclofenac; 2/30 placebo.	0/30 diclofenac; 0/30 placebo.	0/30 diclofenac; 0/30 placebo.	3
Ginsberg & Famaey, 1991	Indomethacin, 4% spray; placebo, spray	Tendinitis	n = 30 crossover 2 x 2 weeks	2–4 sprays 3–5 times daily lightly massaged into skin	1. VAS 2. 4-point verbal rating	1. Indomethacin better than placebo on various pain indices. 2. Subjective improvement: 26/30 indomethacin; 18/30 placebo.	2/30 indomethacin; 0/30 placebo.	0/30 indomethacin; 0/30 placebo.	2
Gui, *et al.*, 1982	Ibuprofen, cream; placebo, cream	Osteoarthritis	n = 40 parallel group 21 days	Application twice daily	Spontaneous pain and pain on pressure and movement	1. Improved spontaneous pain: 17/19 ibuprofen; 9/20 placebo. 2. Improved pain on movement: 14/19 ibuprofen; 7/20 placebo.	0/19 ibuprofen; 0/20 placebo.	0/19 ibuprofen; 0/20 placebo.	3
Hohmeister, 1983	Flufenamate, 3%, plus salicylate, 2%, gel; placebo, gel	Cervical and lumbar back pain	n = 100 parallel group 7, 14, 21 days	Three times daily	Symptom improvement, complete PR	Complete PR at 21 days: 38/49 active gel; 3/51 placebo.	8/49 active gel; 0/51 placebo.	Withdrawals: 0/49 active gel; 0/51 placebo.	4
Mattara, *et al.*, 1994	Flurbiprofen, 40 mg patch; placebo, patch	Scapulo-humoral periarthritis	n = 80 parallel group 14 days	Twice daily	VAS PI for extension, flexion and abduction	Day 14 no pain or slight pain: 14/40 flurbiprofen; 13/40 placebo.	4/40 flurbiprofen; 1/40 placebo.	Withdrawals: 0/40 flurbiprofen; 0/40 placebo. Adverse effects: 5/40 flurbiprofen; 2/40 placebo.	4
Rose, *et al.*, 1991	Piroxicam, 0.5% gel; placebo, gel	Gonarthrosis	n = 30 parallel group 14 days	1 g gel, four times daily	1. Pain on movement 2. Pain at rest 3. Patient global	1. No pain 7/15 piroxicam; 2/15 placebo. 2. Excellent/good: 8/15 piroxicam; 5/15 placebo.	1/15 piroxicam; 1/15 placebo.	0/15 piroxicam; 0/15 placebo.	2
Roth, 1995	Diclofenac, gel; placebo, gel	Osteoarthritis breakthrough pain	n = 119 parallel group 14 days	Four times daily for 2 weeks	Overall pain	NSD.	12/59 diclofenac; 26/60 placebo.	Adverse effects withdrawals: 3/59 diclofenac; 4/60 placebo.	4

TABLE 37 *Active controlled trials in chronic painful conditions: trial design, outcome measures and results*

Study	Drug(s)	Condition	Numbers, study design and follow-up	Dosing regimen	Outcome measures	Analgesic outcome results	Skin irritation	Withdrawals and adverse effects	Quality score
Ammer, 1991	Diclofenac, gel; indomethacin, 1% gel	Soft tissue rheumatism with pain of medium intensity	n = 227 parallel group 14 days	2–4 days per week	1. Pain at rest and on movement 2. General efficacy	1. NSD. 2. Good/excellent: 76/89 diclofenac; 62/84 indomethacin.		Adverse effects withdrawal: 1/89 diclofenac; 0/84 indomethacin.	2
Balthazar-Letawe, 1987	Diclofenac, gel; indomethacin, gel	Rheumatological disorders	n = 50 parallel group 7, 14 days	Twice daily	1. Symptom intensity 3-point scale 2. Investigator global	Improved at 14 days: 15/25 diclofenac; 17/25 indomethacin.	0/25 diclofenac; 0/25 indomethacin.	Withdrawals and adverse effects: 0/25 diclofenac; 0/25 indomethacin.	4
Browning & Johson, 1994	Normal oral NSAID; half normal oral plus piroxicam	Mild to moderate osteoarthritis	n = 191 parallel group, open study 14, 28 days	Three to four times daily, piroxicam	1. Patients' assessment of pain and stiffness 2. Patients' overall assessment of efficacy day	1. Significant reduction in mean score for tenderness and restiction of active movement for topical NSAID. 2. Patients' overall assessment of efficacy day, excellent/good: 54/85 oral alone; 71/106 oral plus topical NSAID.	1/106 piroxicam.	Withdrawal: 1/106 piroxicam. Adverse effects: 1/106 piroxicam; 1/85 oral alone.	2
Dickson, 1991	Piroxicam, 0.5% gel; oral ibuprofen	Chronic osteo-arthritis of knee	n = 235 multicentre, parallel group, double-dummy 4 weeks	1 g gel three times daily; 400 mg ibuprofen three times daily	1. Pain 9-point scale 2. Analgesic consumption 3. Global 4-point scale	NSD between treatments; patient global rating good/better: 68/117 piroxicam; 65/118 ibuprofen.	3/117 piroxicam; 4/118 ibuprofen.	Adverse effects: 30/117 piroxicam; 27/118 ibuprofen. Withdrawals: 9/117 piroxicam; 7/118 ibuprofen.	4
Geller, 1980	Diethylamine salicylate, 10% gel; etofenamate, 5% gel	Chronic disorders	n = 50 crossover 7 days; 4-day washout	Not recorded	1. Pain at rest and in movement, 4-point scale 2. Global 5-point scale patient	1. Diethylamine salicylate better than etofenamate on all scores. 2. After first phase, good/very good results patient global: 24/25 salicylate; 8/25 etofenamate.	Two local effects but drug responsible not given.	Not reported.	2
Giacovazzo, 1992	Diclofenac, gel; felbinac, gel (biphenyl acetic acid)	Osteoarthritis	n = 40 parallel group 1 week	Diclofenac, 160 mg/day, felbinac, 90 mg/day, three times daily	VAS PI	No difference between two treatments. Improvement in pain scores: 14/20 diclofenac; 14/20 felbinac.	0/20 diclofenac; 0/20 felbinac.	0/20 diclofenac; 0/20 felbinac.	1
Golden, 1978	Triethylamine salicylate, 10% cream; oral aspirin, 325 mg tablet	Rheumatic pain	n = 40 parallel group, double-dummy 7 days	Application of cream four times daily; two tablets four times daily	Daily diaries, categorical scales	Good/excellent results: 13/20 salicylate cream; 10/20 oral aspirin.	1/20 salicylate cream; 1/20 oral aspirin.	2/20 salicylate cream; 6/20 oral aspirin.	3
Matucci-Cerinic & Casini, 1988	Ketoprofen, 2.5% gel; etofenemate, 5% gel	Soft tissue rheumatic disorders	n = 36 parallel group 3, 7 days	Twice daily	VAS PI and tenderness	Ketoprofen better than etofenemate for pain on active and passive movement.	0/18 ketoprofen; 0/18 etofenemate.	0/18 ketoprofen; 0/18 etofenemate.	2
Reginster, *et al.*, 1990	Indomethacin, 1% gel; indomethacin, 4% spray	Rheumatoid arthritis	n = 20 crossover 14 days	Three times daily, 100 mg daily total	% improvement on swelling and pain at rest on flexion	Both improved significantly from baseline.	2/20; 2/20.	Withdrawals and adverse effects: 0/20; 0/20.	2
Ritchie, 1996	Flurbiprofen, patch; piroxicam, 0.5% gel	Soft tissue rheumatism of shoulder or elbow	n = 131 crossover at 4 days 4, 8, 14 days	Flurbiprofen, 40 mg patch, twice daily; 3 cm piroxicam gel four times daily	Pain, tenderness	Statistically more PR with flurbiprofen.		Adverse effects withdrawals: 1/133 flubiprofen; 3/133 piroxicam.	3
Rosenthal & Bahous, 1993	DHEP, 1% plaster; diclofenac, 1% gel	Periarticular, tendinous inflammations	n = 190 parallel group 14 days	Plaster twice daily; gel four times daily	Spontaneous pain, pain on pressure, patient global	Patient global, good/excellent: 79/96 plaster; 39/94 gel.	2/96 plaster; 3/94 gel.	Withdrawals and adverse effects: 0/96 plaster; 0/94 gel.	3
Vitali, 1980	Ketoprofen, 1, 2.5 and 5% gel	Orthopaedic	n = 62 parallel group 14 days	5–15 cm twice daily for 6–13 days	Spontaneous pain, palpation, movement	1. 2.5% gel was most useful. 2. Spontaneous pain better/much better: 13/20 1% gel; 18/20 2.5% gel; 16/20 5% gel.	Not reported.	Not reported.	3

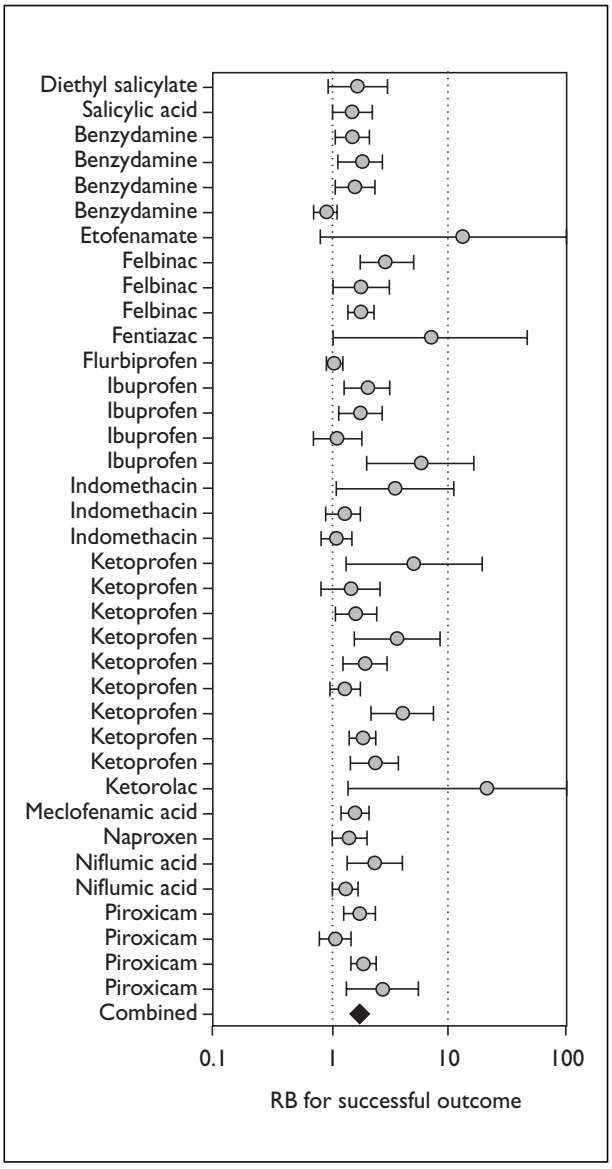


FIGURE 33 *Placebo-controlled trials of topical NSAIDs: one week outcome in acute painful conditions*

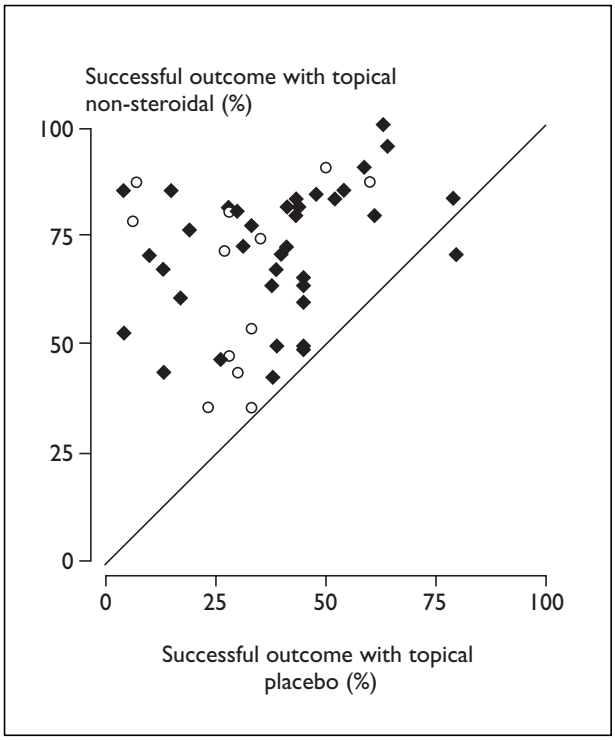


FIGURE 34 *Placebo-controlled trials of topical NSAIDs in acute (◆) and chronic (○) conditions*

Relative benefit and 95% CIs for each drug compared with placebo are shown in *Figure 36*. Seven of the 12 studies showed statistical superiority for topical NSAIDs over placebo. The scatter of the proportion of patients with at least 50% pain relief with topical NSAID or placebo is shown in *Figure 34*. All 12 comparisons were in the segment favouring treatment over placebo. The one trial which did not have dichotomous outcomes also reported statistical benefit for topical NSAIDs over placebo.

The pooled relative benefit for all 12 comparisons was 2.0 (95% CI, 1.5–2.7) and the NNT was 3.1 (95% CI, 2.7–3.8) *(Table 38)*. Sensitivity analysis by quality score or treatment group size produced no significant change in these estimates. No single topical NSAID was tested in as many as three placebo-controlled studies and combined estimates could not therefore be calculated for any single drug.

The percentage of patients achieving at least 50% pain relief with active drug treatment or placebo in all studies in all trials (placebo and active drug controlled) in chronic conditions is shown in *Figure 35* (upper panel). The range with placebo was from 5% to 60%. With topical NSAIDs it was from 30% to 95%. There was no significant difference in the (low) frequency of local or systemic adverse effects, or drug-related withdrawal *(Table 38)*.

Comparison with oral NSAIDs

Five studies compared topical with oral NSAIDs, three in acute[11,12,13] and two in chronic conditions.[14,15] None showed statistical benefit of oral over topical NSAIDs.

Comments

Topical NSAIDs were significantly more effective than placebo. This is not just due to rubbing. Placebo preparations were also rubbed on affected parts. The significant difference was therefore additional to any effect of rubbing. Topical pre-

TABLE 38 *Combined results and sensitivity analysis for topical NSAIDs in acute and chronic painful conditions*

Drug	Trials	Patients	Average number of treated patients	CER	EER	RR	NNT
Acute painful conditions							
Combined efficacy data	37	3239	47	39	71	1.7 (1.5–1.9)	3.9 (3.4–4.4)
Local adverse effects				3	2.6	1.2 (0.8–1.7)	
Systemic adverse effects				0.7	0.8	1.0 (0.6–1.8)	
Withdrawal due to adverse effects				0.4	0.6	0.8 (0.4–1.4)	
Trials with quality score 3–5 only	30	2834	52	38	72	1.7 (1.5–1.9)	3.9 (3.4–4.4)
Trials with treatment groups of < 40 patients	20	933	24	35	76	1.9 (1.6–2.2)	2.6 (2.3–3.1)
Trials with treatment groups of 40–80 patients	8	810	51	44	66	1.6 (1.1–2.2)	5.0 (3.7–7.4)
Trials with treatment groups of > 80 patients	7	1496	123	41	67	1.6 (1.3–1.9)	4.6 (3.7–5.9)
Ketoprofen	9	724	43	36	74	2.0 (1.5–2.6)	2.6 (2.3–3.2)
Felbinac	3	413	70	32	66	2.0 (1.5–2.7)	3.0 (2.4–4.1)
Ibuprofen	4	284	36	34	70	1.9 (1.2–3.0)	3.5 (2.5–5.6)
Piroxicam	4	589	74	39	69	1.6 (1.2–2.2)	4.2 (3.1–6.1)
Benzydamine	4	245	31	62	84	1.4 (0.9–2.0)	6.7 (3.8–23)
Indomethacin	3	394	66	32	47	1.3 (0.9–1.8)	10 (5.0–∞)
Chronic painful conditions							
Combined efficacy data	12	1097		30	65	2.0 (1.5–2.7)	3.1 (2.7–3.8)
Local adverse effects				5.3	5.9	0.9 (0.4–1.7)	
Systemic adverse effects				1.3	1.1	1.1 (0.5–2.3)	
Withdrawal due to adverse effects				0.7	0.7	1.0 (0.4–2.4)	
Trials with quality score 3–5 only	9	987	55	27	62	2.2 (1.5–3.1)	3.1 (2.6–3.8)
Trials with treatment groups of < 40 patients	6	261	22	31	69	2.2 (1.5–3.1)	2.6 (2.0–3.6)
Trials with treatment groups of > 40 patients	6	836	70	29	61	2.0 (1.7–2.4)	3.3 (2.8–4.3)

Response is either the proportion of patients with successful outcome or percentage of patients with an adverse effect. An infinite NNT CI indicates that there may be no benefit from the treatment compared with placebo.

parations produced NNTs in the range 3–5 *(Table 38)*. At least one patient in about three using a topical NSAID will achieve at least 50% pain relief who would not have done had they used placebo.

While this result may surprise some, it is not because the trials were poor. Placebo-controlled studies in both acute and chronic conditions had quality scores of 3 or more on a scale of 1–5 in over 75% of reports (see *Table 33)*. This is important, since trials of lower methodological quality (2 or less using the same validated scale as here) have been shown to have a more favourable outcome.[16]

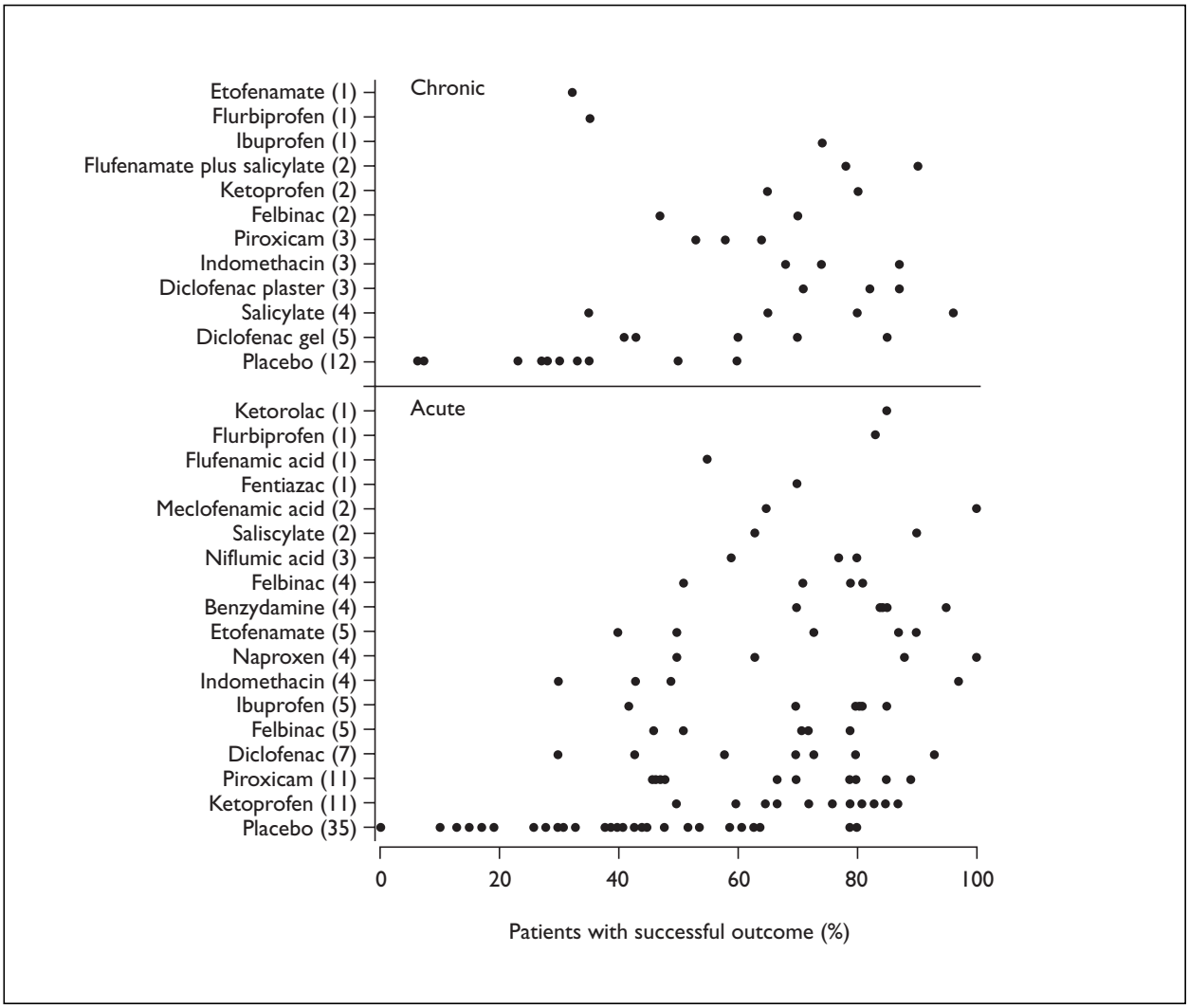


FIGURE 35 *Successful outcome for treatment arms from placebo- and active controlled trials in chronic and acute painful conditions – drug and number of trials*

It was judged sensible to pool only data for individual drugs when there were at least three RCTs. In acute conditions there was enough information to make comparisons *(Table 38)*. The average response for placebo was similar for individual drugs apart from benzydamine. Ketoprofen, felbinac, ibuprofen and piroxicam were all statistically superior to placebo, in contrast to indomethacin and benzydamine which were not. CIs for the NNT for ketoprofen did not overlap with those of benzydamine or indomethacin. There is no clear message as to which of ketoprofen, felbinac, ibuprofen or piroxicam was best, or indeed whether there was any difference in efficacy. They all work.

Local skin reactions were rare (3.6%) and systemic effects were rarer (less than 0.5%). Local or systemic adverse effects of sufficient severity to cause withdrawal from the study were also rare (0.5%). Adverse effects were no more common than with placebo.

Topical NSAIDs are less associated with the gastrointestinal adverse effects seen with the same drugs taken orally.[17] The low incidence of systemic adverse effects for topical NSAIDs probably results from the much lower plasma concentrations from similar doses applied topically to those administered orally.[13,18] Topical application of ibuprofen resulted in significant tissue concentrations in deep tissue compartments, more than enough to inhibit inflammatory enzymes.[18,19]

These positive results for topical NSAIDs could, it may be argued, be skewed by publication restricted to positive findings. It is next to impossible to rebut this argument. Strenuous efforts were made to

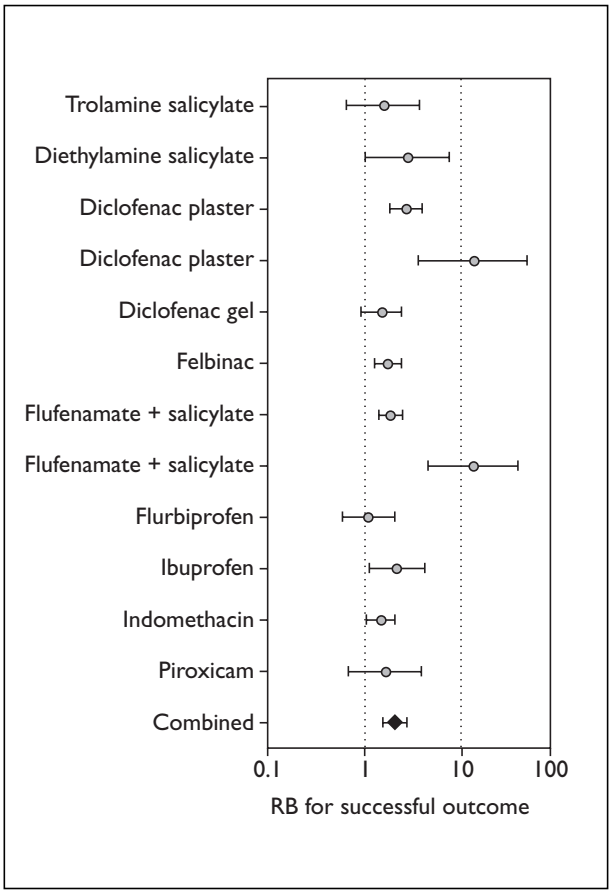


FIGURE 36 *Placebo-controlled trials of topical NSAIDs: 2-week outcome in chronic painful conditions*

unearth unpublished data. Ironically, one pharmaceutical company withheld results they claimed to be positive and favourable to their product. Rosenthal's file drawer argument[20] says there would need to be many negative results (more than 692 for acute, 37 for chronic) to overturn these positive results.

More important is the empirical evidence that small trials (arbitrarily set at fewer than 40 patients per group as being between the mean and median sizes of 47 and 32 patients per treated group) produced exaggerated estimates of clinical efficacy by 24% (4.8 minus 3.9/3.9, *Table 38)* with CIs which did not overlap. By contrast, trial quality made no difference despite evidence to the contrary from other settings.[16] Size of treatment group may be an important issue for credibility of estimates of clinical efficacy in treatments, just like randomisation[21,22] and double-blinding.[21] Just as it may be hazardous to change practice on the basis of a single small trial, similarly beware meta-analysis restricted to multiple small trials.[23]

The important research agenda is to identify those patients with chronic disease, particularly elderly patients, who may benefit from using topical rather than oral NSAIDs. We need to compare the pain relief and mobility, harm and cost for these alternatives. The few studies identified which compared oral with topical NSAIDs were of inadequate design and power to answer these important questions. In the meantime, the message is that topical NSAIDs are effective and safe.

References

1. Wynne HA, Campbell M. Pharmacoeconomics of nonsteroidal anti-inflammatory drugs (NSAIDs). *PharmacoEconomics* 1993;**3**:107–23.
2. Bateman DN, Kennedy JG. Non-steroidal anti-inflammatory drugs and elderly patients. The medicine may be worse than the disease. *BMJ* 1995;**310**:817–18.
3. Rational use of NSAIDs for musculoskeletal disorders. *Drug Therapeut Bull* 1994;**32**:91–5.
4. Jadad AR, Carroll D, Moore A, McQuay H. Developing a database of published reports of randomised clinical trials in pain research. *Pain* 1996;**66**:239–46.
5. Jadad AR, Moore RA, Carroll D, Jenkinson C, Reynolds DJM, Gavaghan DJ, *et al.* Assessing the quality of reports of randomized clinical trials: is blinding necessary? *Control Clin Trials* 1996;**17**:1–12.
6. L'Abbé KA, Detsky AS, O'Rourke K. Meta-analysis in clinical research. *Ann Intern Med* 1987;**107**:224–33.
7. DerSimonian R, Laird N. Meta-analysis of clinical trials. *Control Clin Trials* 1986;**7**:177–88.
8. Yusuf S, Peto R, Lewis J, Collins R, Sleight P. Beta-blockade during and after myocardial infarction: an overview of the randomized trials. *Prog Cardiovasc Dis* 1985;**27**:335–71.
9. Laupacis A, Sackett DL, Roberts RS. An assessment of clinically useful measures of the consequences of treatment. *N Engl J Med* 1988;**318**:1728–33.
10. Cook RJ, Sackett DL. The number needed to treat: a clinically useful measure of treatment effect. *BMJ* 1995;**310**:452–4.
11. Åkermark C, Forsskåhl B. Topical indomethacin in overuse injuries in athletes. A randomized double blind study comparing Elmetacin® with oral indomethacin and placebo. *Int J Sports Med* 1990;**11**:393–6.
12. Hosie GAC. The topical NSAID, felbinac, versus oral ibuprofen: a comparison of efficacy in the treatment of acute lower back injury. *Br J Clin Res* 1993;**4**:5–17.
13. Vanderstraeten G, Schuermans P. Study on the effect of etofenamate 10% cream in comparison with an oral NSAID in strains and sprains due to sports injuries. *Acta Belg Med Phys* 1990;**13**:139–41.

14. Browning RC, Johson K. Reducing the dose of oral NSAIDs by use of feldene gel: an open study in elderly patients with osteoarthritis. *Adv Ther* 1994;**11**:198–206.

15. Golden EL. A double-blind comparison of orally ingested aspirin and a topically applied salicylate cream in the relief of rheumatic pain. *Curr Ther Res* 1978;**24**:524–9.

16. Khan KS, Daya S, Jadad AR. The importance of quality of primary studies in producing unbiased systematic reviews. *Arch Intern Med* 1996;**156**:661–6.

17. Evans JMM, McMahon A, McGilchrist M, White G, Murray F, McDevitt D, *et al.* Topical non-steroidal anti-inflammatory drugs and admission to hospital for upper gastrointestinal bleeding and perforation: a record linkage case-control study. *BMJ* 1995;**311**:22–6.

18. Berner G, Engels B, Vögtle-Junkert U. Percutaneous ibuprofen therapy with Trauma-Dolgit gel: bioequivalence studies. *Drugs Exp Clin Res* 1989;**15**:559–64.

19. Treffel P, Gabard B. Feasibility of measuring the biavailability of topical ibuprofen in commercial formulations using drug content in epidermis and a methyl nicotinate skin inflammation assay. *Skin Pharmacol* 1993;**6**:268–75.

20. Rosenthal R. The "file drawer problem" and tolerance for null results. *Psychol Bull* 1979;**86**:638–41.

21. Schulz KF, Chalmers I, Hayes RJ, Altman DG. Empirical evidence of bias: dimensions of methodological quality associated with estimates of treatment effects in controlled trials. *JAMA* 1995;**273**:408–12.

22. Carroll D, Tramèr M, McQuay H, Nye B, Moore A. Randomization is important in studies with pain outcomes: systematic review of transcutaneous electrical nerve stimulation in acute postoperative pain. *Br J Anaesth* 1996;**77**:798–803.

23. Egger M, Smith GD. Misleading meta-analysis. Lessons from "an effective, safe, simple" intervention that wasn't. *BMJ* 1995;**310**:752–4.

Studies included

Airaksinen O, Venäläinen J, Pietiläinen T, 1993. Ketoprofen 2.5% gel versus placebo gel in the treatment of acute soft tissue injuries. *Int J Clin Pharmacol Ther Toxicol*;**31**:561–3.

Åkermark C, Forsskåhl B, 1990. Topical indomethacin in overuse injuries in athletes. A randomized double blind study comparing Elmetacin® with oral indomethacin and placebo. *Int J Sports Med*;**11**:393–6.

Algozzine GJ, Stein GH, Doering PL, Araujo OE, Akin KC, 1982. Trolamine salicylate cream in osteoarthritis of the knee. *JAMA*;**247**:1311–13.

Ammer K, 1991. Perkutane antirheumatische therapie beim weichteilrheumatismus. *Rheumatologie*;**6**:412–21.

Aoki T, *et al.*, 1984. Comparative study of piroxicam gel with indomethacin gel and piroxicam placebo gel in the treatment of pain from orthopedic trauma. *Jpn Pharmacol Ther*;**12**:101–17.

Arioli G, Scaramelli M, Pillosu W, 1990. La terapia topica delle affezioni acte musculoscheletriche. Risultati di uno studio comparativo sul piroxicam crema 1% versus diclofenac emulgel 1% [Topical therapy of acute skeletal muscle diseases. Results of a comparative study on piroxicam cream 1% versus diclofenac emulgel 1%]. *Clin Ter*;**134**:363–9.

Auclair J, Georges M, Grapton X, Gryp L, d'Hooghe M, Meisser RG, *et al.*, 1989. A double-blind controlled multicenter study of percutaneous niflumic acid gel and placebo in the treatment of achilles heel tendinitis. *Curr Ther Res*;**46**:782–8.

Baixauli F, Inglés F, Alcántara P, Navarrete R, Puchol E, Vidal F, 1990. Percutaneous treatment of acute soft tissue lesions with naproxen gel and ketoprofen gel. *J Int Med Res*;**18**:372–8.

Balthazar-Letawe D, 1987. Voltaren Emulgel en practique rheumatologique. Essai comparatif avec Indocid gel [Voltaren Emulgel in rheumatological practice. A comparative trial with Indocid gel]. *Acta Belg Med Phys*;**10**:109–10.

Baracchi G, Messina Denaro S, Piscini S, 1982. Esperienza sull'impiego topico dell'acido isobutil-fenil-proprioico (Ibuprofen) nell flogosi traumatica. Confronto a doppia cecità con placebo [Experience of the topical use of isobutylfenylproprionic acid (Ibuprofen) in traumatic inflammation. A double-blind comparison with placebo]. *Gazz Med Ital*;**141**:691–4.

Bolten W, 1991. Felbinac-Gel zur Behandlung lokalisierter extraartikulärer rheumatischer Beschwerden – eine multizentrische, placebokontrollierte, randomisierte Studie [Felbinac gel for treatment of localized extra-articular rheumatic diseases – a multicentre, placebo controlled, randomized study]. *Z Rheumatol*;**50**:10–13.

Bouchier-Hayes TA, Rotman H, Darekar BS, 1990. Comparison of the efficacy and tolerability of diclofenac gel (Voltarol Emulgel) and felbinac gel (Traxam) in the treatment of soft tissue injuries. *Br J Clin Pract*;**44**:319–20.

Boutrón F, Galicia A, Zamora G, Martinez-Zurita F, 1994. Single-blind study to evaluate the efficacy and safety of naproxen gel compared with diclophenac emulgel in the treatment of soft tissue injuries. *Proc West Pharmacol Soc*;**37**:153–6.

Browning RC, Johson K, 1994. Reducing the dose of oral NSAIDs by use of feldene gel: an open study in elderly patients with osteoarthritis. *Adv Ther*;**11**:198–206.

Campbell J, Dunn T, 1994. Evaluation of topical ibuprofen cream in the treatment of acute ankle sprains. *J Accid Emerg Med*;**11**:178–82.

Camus J, 1975. Action de la Myrtécaïne associée au Salicylate de Diéthylamine, en traitement local, dans diverses affections rheumatismales. *Rheumatologie*;**27**:61–6.

Candela V, Bagarone A, 1986. Studio in cieco semplice sull'attivita terapeutica del "Fastum Gel" nelle lesioni traumatiche da sport. *Med Sport*;**39**:57–63.

Chatterjee DS, 1977. A double-blind clinical study with benzydamine 3% cream on soft tissue injuries in an occupational health centre. *J Int Med Res*;**5**:450–8.

Commandre F, Zakarian H, Corriol-Rohou S, 1993. Comparison of the analgesic and anti-inflammatory effects of topical niflumic acid gel versus piroxicam gel in the treatment of musculoskeletal disorders. *Curr Ther Res*;**53**:113–21.

Curioni GB, Di Domenica F, Daolio P, Spignoli G, 1985. Valutazione dell'efficacia terapeutica di ibudros gel in traumatologia. *Clin Europa*;**24**:456–60.

Dickson DJ, 1991. A double-blind evaluation of topical piroxicam gel with oral ibuprofen in osteoarthritis of the knee. *Curr Ther Res*;**49**:199–207.

Diebschlag W, 1986. Diclofenac bei stumpf-traumatischen Sprunggelenkschwellungen. Volumetrische Verlaufsbestimmung im plazebokontrollierten Doppelblindversuch [Diclofenac in blunt traumatic ankle joint swelling. Volumetric monitoring in a placebo controlled double-blind trial]. *Fortschr Med*;**104**:437–40.

Diebschlag W, Nocker W, 1987. Einfluß einer topischen Behandlung auf den Krankheitsverlauf bei Sprunggelenks-Distorsionen [The effect of topical treatment in the treatment of disease in ankle sprains]. *Arzneimittelforschung*;**37**:1076–81.

Diebschlag W, Nocker W, Bullingham R, 1990. A double-blind study of the efficacy of topical ketorolac tromethamine gel in the treatment of ankle sprain, in comparison to placebo and etofenamate. *J Clin Pharmacol*;**30**:82–9.

Diebschlag W, Nocker W, Lehmacher W, 1992. Treatment of acute sprains of the ankle joint. A comparison of the effectiveness and tolerance of two gel preparations containing indomethacin. *Fortschr Med*;**110**:64–72.

Dreiser RL, 1988. Clinical trial of efficacy and tolerability of topical ibuprofen in the treatment of tendinitis. *J Int Méd*;**119**:15–31.

Dreiser RL, 1989. Clinical trial – Fastum gel FG-6. Unpublished.

Dreiser RL, Tisne-Camus M, 1993. DHEP plasters as a topical treatment of knee osteoarthritis – a double-blind placebo-controlled study. *Drugs Exptl Clin Res*;**19**:117–23.

Dreiser RE, Charlot J, Lopez A, Ditisheim A, 1990. Clinical evaluation of niflumic acid gel in the treatment of uncomplicated ankle sprains. *Curr Med Res*;**12**:93–9.

Dreiser RL, Roche R, de-Sahbe R, Thomas F, Leuteneger E, 1994. Flurbiprofen local action transcutaneous: clinical evaluation in the treatment of acute ankle sprains. *Eur J Rheum Inflamm*;**14**:9–13.

El-Hadidi T, El-Garf A, 1991. Double-blind study comparing the use of Voltaren Emulgel versus regular gel during ultrasonic sessions in the treatment of localized traumatic and rheumatic painful conditions. *J Int Med Res*;**19**:219–27.

Fantato S, de Gregorio M, 1971. Clinical evaluation of topical benzydamine in traumatology. *Arzneimittelforschung*;**21**:1530–5.

Fotiades P, Bach GL, 1976. Wirkung einer flufenaminsäurehaltigen salbe bei verschiedenen rheumatischen erkrankungen. *Fortschr Med*;**94**:1036–8.

Frahm E, Elsasser U, Kämmereit A, 1993. Topical treatment of acute sprains. *Br J Clin Pract*;**47**:321–2.

Fujimaki E, *et al.*, 1985. Clinical evaluation of piroxicam gel versus indomethacin gel and placebo in the treatment of muscle pain: a double-blind, multicenter study. *Jpn Pharmacol Ther*;**12**:119–37.

Galeazzi M, Marcolongo R, 1993. A placebo-controlled study of the efficacy and tolerability of a nonsteroidal anti-inflammatory drug, DHEP plaster, in inflammatory peri- and extra-articular rheumatological diseases. *Drugs Exp Clin Res*;**19**:107–15.

Gallacchi G, Mautone G, Lualdi P, 1990. Painful inflammatory conditions. Topical treatment with diclofenac hydroxyethylpyrrolidine. *Clin Trials J*;**27**:58–64.

Geller O, 1980. Vergleich eines Salizylat/Heparin-Gels mit einem Mono-Substanzpräparat. Ergebnisse einer Doppelblind-cross-over-Studie [Comparison of a salicylate heparin gel with a monosubstance preparation. Results of a double-blind crossover study]. *Münch Med Wschr*;**122**:1231–2.

Giacovazzo M, 1992. Clinical evaluation of a new NSAID applied topically (BPAA Gel) vs. diclofenac emulgel in elderly osteoarthritic patients. *Drugs Exp Clin Res*;**18**:201–3.

Ginsberg F, Famaey JP, 1991. Double-blind, randomized crossover study of the percutaneous efficacy and tolerability of a topical indomethacin spray versus placebo in the treatment of tendinitis. *J Int Med Res*;**19**:131–6.

Golden EL, 1978. A double-blind comparison of orally ingested aspirin and a topically applied salicylate cream in the relief of rheumatic pain. *Curr Ther Res*;**24**:524–9.

Governali E, Casalini D, 1995. Ricerda clinica controllata tra ketoprofene gel 5% e ketoprofene crema 1% in pazienti con postumi di lesioni traumatiche. *Riabilitazione*;**28**:61–9.

Gualdi A, Bonollo L, Martini A, Forgione A, 1987. Antinflammatori no steroidi per uso topico in traumatologia: studio clinico con flunoxaprofene e chetoprofene. *Riforma Med*;**102**:401–4.

Gui L, Pellacci F, Ghirardini G, 1982. Impiego dell'ibuprofen crema in pazienti ambulatoriali di interesse ortopedico. Confronto in doppia cecità con placebo. *Clin Ter*;**101**:363–9.

Haig G, 1986. Portable thermogram technique for topically applied benzydamine cream in acute soft-tissue injuries. *Int J Tissue React*;**8**:145–7.

Hallmeier B, 1988. Efficacy and tolerance of etofenamate and diclofenac in acute sports injuries. *Rheuma*;**8**:183–6.

Hallmeier B, Michelbach B, 1986. Etofenamat unter tapeverbänden. *Med Welt*;**37**:1344–8.

Hohmeister R, 1983. Die behandlung von weichteilrheumatischen erkrankungen mit mobilisin gel. *Fortschr Med*;**101**:1586–8.

Hosie GAC, 1993. The topical NSAID, felbinac, versus oral ibuprofen: a comparison of efficacy in the treatment of acute lower back injury. *Br J Clin Res*;**4**:5–17.

Julien D, 1989. Clinical trial – Fastum gel FG-8. Unpublished.

Kockelbergh M, Verspeelt P, Caloine R, Dermaux F, 1985. Traitement anti-inflammatoire local par un gel de kétoprofène: données cliniques récentes [Local anti inflammatory treatment with a ketoprofen gel: current clinical findings]. *Acta Belg Med Phys*;**8**:205–13.

Kroll MP, Wiseman RL, Guttadauria M, 1989. A clinical eveluation of piroxicam gel: an open comparative trial with diclofenac gel in the treatment of acute musculoskeletal disorders. *Clin Ther*;**11**:382–91.

Lester AA, 1983. Management of sprained ankles. *Practitioner*;**225**:935–6.

Linde F, Hvass I, Jürgensen U, Madsen F, 1985. Ankelforstuvninger behandlet med benzydamin 5% creme. *Ugeskr Laeger*;**148**:12–13.

Mattara L, Trotta F, Biasi D, Cervetti R, 1994. Evaluation of the efficacy and tolerability of a new locally acting preparation of flurbiprofen in scapulohumeral periarthritis. *Eur J Rheum Inflamm*;**14**:15–20.

Matucci-Cerinic M, Casini A, 1988. Ketoprofen vs. etofenamate in a controlled double-blind study: evidence of topical effectiveness in soft tissue rheumatic pain. *Int J Clin Pharmacol Res*;**8**:157–60.

McLatchie GR, McDonald M, Lawrence GF, Rogmans D, Lisai P, Hibberd M, 1989. Soft tissue trauma: a randomised controlled trial of the topical application of felbinac, a new NSAID. *Br J Clin Pract*;**43**:277–80.

Montagna CG, Turroni L, Martinelli D, Orlandini MC, 1990. Single-blind comparative study of meclofenamic acid gel versus naproxen gel in acute musculoskeletal disorders. *Curr Ther Res*;**47**:933–9.

Morris WD, Scott HV, Peters WA, Ketelbey JW, 1991. Felbinac topical gel for acute soft tissue sports injuries. *N Z J Sports Med*;**19**:45–7.

Noret A, Roty V, Allington N, Hauters P, Zuinen C, Poels R, 1987. Ketoprofen gel as topical treatment for sport injuries. *Acta Ther*;**13**:367–78.

Oakland C, 1993. A comparison of the efficacy of the topical NSAID felbinac and ultrasound in the treatment of acute ankle injuries. *Br J Clin Res*;**4**:89–96.

Parrini M, Cabitza P, Arrigo A, Vanasia M, 1992. Efficacia e tollerabilità del ketoprofene sale di lisina schiuma per uso topico nel trattamento di alcune patologie traumatiche dell'apparato locomotore [Efficacy and tolerability of ketoprofen lysine salt foam for topical use in the treatment of traumatic pathologies of the locomotor apparatus]. *Clin Ter*;**141**:199–204.

Picchio AA, Volta S, Longoni A, 1981. Studio clinico controllato sull'impiego dell'ibuprofen per uso topico in traumatologica sportiva. *Med Sport*;**34**:403–6.

Pineda MD, Capulong OM, de Guzman DL, Te CV, 1993. The topical NSAIDs, felbinac and piroxicam: a comparison of efficacy and safety in the treatment of acute soft tissue trauma. *Br J Clin Res*;**4**:63–72.

Ramesh N, Steuber U, 1983. Dolgit® cream in accident- and sports-related injuries in medical practice. Results of a double blind study. *Therapiewoche*;**33**:4563–70.

Reginster JY, Crommen J, Renson M, Franchimont P, 1990. Percutaneous administration of indomethacin in rheumatoid arthritis. *Curr Ther Res*;**47**:548–53.

Ritchie LD, 1996. A clinical evaluation of flurbiprofen LAT and piroxicam gel: a multicentre study in general practice. *Clin Rheumatol*;**15**:243–7.

Rose W, Manz G, Lemmel E, 1991. Behandlung der akitivierten gonarthrose mit topisch appliziertum piroxicam-gel. *Münch Med Wschr*;**133**:562–6.

Rosemeyer B, 1991. Behandlung von distorsionen des sprunggelenkes. *Med Welt*;**42**:166–70.

Rosenthal M, Bahous I, 1993. A controlled clinical study on the new topical dosage form of DHEP plasters in patients suffering from localized inflammatory diseases. *Drugs Exp Clin Res*;**19**:99–105.

Russell AL, 1991. Piroxicam 0.5% topical gel compared to placebo in the treatment of acute soft tissue injuries: a double-blind study comparing efficacy and safety. *Clin Invest Med*;**14**:35–43.

Sanguinetti C, 1989. Trattemento con BPAA gel dei traumi dei tessuti moli. *Clin Ter*;**130**:255–8.

Seligra A, Inglés F, 1990. A comparative study of naproxen gel and flufenamic acid gel in the treatment of soft tissue injuries. *Curr Med Res*;**12**:249–54.

Sinniger M, Blanchard P, 1981. Controlled clinical trial with fentiazec cream in sport microtraumatology. *J Int Med Res*;**9**:300–2.

Sugioka Y, *et al.,* 1984. Multicenter clinical evaluation of piroxicam gel vs. indomethacin gel in the treatment of non-traumatic diseases of tendon or muscle. *Jpn Pharmacol Ther*;**12**:139–53.

Taboada A, 1992. Experencia controlada con gel de piroxicam asociado a ultrasonidos en afecciones agudas del aparato locomotor. *Prens Méd Argent*;**79**:630–2.

Thorling J, Linden B, Berg R, Sandahl A, 1990. A double-blind comparison of naproxen gel and placebo in the treatment of soft tissue injuries. *Curr Med Res*;**12**:242–8.

Tonutti A, 1993. Utilizzazione del ketoprfene gel 5% (Oridis gel) nella pratica traumatologica: studio in doppio cieco controllato verso etofenamato. *Ort Traum*;**XIV**:119–25.

Vanderstraeten G, Schuermans P, 1990. Study on the effect of etofenamate 10% cream in comparison with an oral NSAID in strains and sprains due to sports injuries. *Acta Belg Med Phys*;**13**:139–41.

Vecchiet L, Colozzi A, 1991. Effects of meclofenamic acid in the treatment of lesions deriving from minor traumatology. *Clin J Pain*;**7**:S54–9.

Vitali G, 1980. Sperimentazione clinica controllata sull' impiego topico dell'acido 2-(3-benzoil-fenil) proprionico a tre diverse concentrazioni [Controlled clinical experiment with the topical use of 2-(3-benzoyl-phenyl) proprionic acid at 3 different concentrations]. *Clin Ter*;**94**:257–73.

Wanet G, 1979. Etude clinique contrôlée d'un topique associant la nopoxamine au salicylate de diéthylamine (Algésal suractivé) en médecine physique et réhabilitation [Controlled clinical study of a topic associating nopoxamine with diethylamine salicylate (Algésal suractivé) in physical medicine and rehabilitation)]. *J Belge Méd Phys Réhab*;**2**:119–26.

Zerbi E, Pace A, Demarchi F, Bassi F, Arrigo A, Garagiola U, 1992. Ketoprofen lysine salt in a new foam formulation for the topical treatment of traumatic injuries: a controlled, between-patient, clinical trial. *Curr Ther Res*;**51**:823–9.

Chapter 12

Injected morphine in postoperative pain

Summary

The pain relief after injected morphine compared with placebo is examined in this chapter, in patients with moderate or severe pain after surgery, and the efficacy of injected morphine is related to that of oral analgesics. A literature search of various databases was for randomised, single-dose, placebo-controlled trials.

Pain relief or pain intensity difference over 4–6 hours and adverse effects were extracted. The number of patients with at least 50% pain relief was derived and then used to calculate the relative benefit and the NNT for one patient to achieve at least 50% pain relief for 4–6 hours.

In 15 trials intramuscular morphine, 10 mg (486 patients), was compared with placebo (460 patients); the NNT was 2.9 (95% CI, 2.6–3.6). One in three patients with moderate or severe postoperative pain achieved at least 50% pain relief; they would not have done had they been given placebo. Minor adverse effects were more common with morphine (34%) than with placebo (23%) (relative risk 1.49 (95% CI, 1.09–2.04)) but drug-related study withdrawal was rare and not different from placebo. Intramuscular morphine, 10 mg, gives analgesia equivalent to oral NSAID, in keeping with historic results from single trials. For patients who can swallow, oral NSAID may be the best choice.

This chapter of the review has been published in full by McQuay and colleagues.[1]

Introduction

Perhaps understandably, we all tend to believe that injecting drugs provides better pain relief than taking the same drug by mouth. Indeed, it took generations for doctors to be persuaded that oral morphine was effective in cancer pain – they all wanted to inject. This chapter focuses on the post-operative pain relief produced by injection of morphine, using the same methods as for the oral drugs. The aim was to achieve an estimate of the analgesic efficacy of injected morphine which could be compared with the estimates for the oral drugs.

Methods

Single-dose, randomised, placebo-controlled trials of injectable (intramuscular, subcutaneous and intravenous) morphine in acute postoperative pain were sought. A number of different search strategies were used to identify eligible reports in MEDLINE (1966–97), EMBASE (1980–97), the Cochrane Library (1997 issue 2) and the Oxford Pain Relief Database (1950–94).[2] The last electronic search was conducted in March 1997. The words 'morphine', 'diamorphine', 'heroin' were used to identify relevant reports, using a combination of free text words and MeSH terms, and without restriction to language. Additional reports were identified from reference lists of retrieved reports, review articles, and specialist textbooks.

Included and excluded reports

Inclusion criteria were full journal publication of RCTs which included single-dose treatment groups of injected (intravenous, intramuscular or subcutaneous) morphine and placebo, acute postoperative pain, blinded design, baseline pain of moderate to severe intensity, adult patients, and assessments of pain intensity or pain relief over 4–6 hours with results for TOTPAR, SPID, VAS TOTPAR, or VAS SPID), or with data from which these could be calculated. Review articles, letters or abstracts were not included.

Reports were screened to eliminate those without pain outcomes, those which were definitely not randomised, or were abstracts or reviews. Each report which could possibly be described as an RCT was read independently by each of the authors and scored using a three-item, 1–5 score, quality scale.[3] Consensus was then achieved. The maximum score for an included study was 5 and the minimum 1.

Data extraction and analysis

Data extracted from the reports were the pain setting, study treatment groups, numbers of patients treated, study duration, the route and dose of morphine, and mean or derived TOTPAR, SPID, VAS TOTPAR or VAS SPID or any dichotomous global pain relief outcome.

Information on minor and major adverse events, as defined by the authors of the original reports, was also extracted.

For each report with mean TOTPAR, SPID, VAS TOTPAR or VAS SPID values for morphine and placebo, the data was converted to percentage of maximum by division into the calculated maximum value.[4] The proportion of patients in each treatment group who achieved at least 50% max-TOTPAR was calculated using verified equations.[5–7] These proportions were then converted into the number of patients achieving at least 50% max-TOTPAR by multiplying by the total number of patients in the treatment group.[5]

Information on the number of patients with > 50% maxTOTPAR for morphine and placebo was used to calculate relative risk (or benefit) and NNT by pooling data when available from at least three comparisons between morphine and placebo with a particular dose and route of administration. Relative risk or benefit estimates were calculated with their 95% CIs using a random effects model[8] for analgesic data which were not homogenous ($p < 0.1$) and a fixed effects model[9] for adverse effect data which were homogenous ($p > 0.1$). Homogeneity of the analgesic results was also explored graphically.[10] The NNT[11] was calculated with a 95% CI.[12] A statistically significant difference from control was assumed when the 95% CIs of the relative risk/benefit did not include 1. Statistical difference between NNTs was assumed when CIs did not overlap.

Results

In all, 18 reports of 20 trials fulfilled the inclusion criteria; 696 patients were given morphine and 563 placebo. No trials of subcutaneous morphine or of diamorphine by any route of administration met the inclusion criteria. Morphine was given by intramuscular injection in all studies except one,[13] in which it was given intravenously. Morphine doses were 5 mg,[14,15] 8 mg,[13,16] 10 mg,[14,15,17,18–29] 12.5 mg[30] and 20 mg.[28] Details of these studies are presented in *Table 39*.

Two studies[18,30] included a mixed population of patients with postoperative and other acute pain. Trials otherwise investigated pain relief predominantly after orthopaedic and gynaecological surgery. Pain outcomes were over 6 hours except for two studies in which they were over 4 hours.[22,30] Quality scores were 2 for two reports, 3 for six, 4 for nine and 5 for one.

Nine reports which appeared to fulfill inclusion criteria were omitted. Three studies[31–33] had pain relief or intensity information for 1 hour or less. Two reports[34,35] appeared to duplicate previously published information and four[36–39] used non-standard assessments which could not be used.

Only for 10 mg doses of intramuscular morphine was data available from at least three trials, which could be pooled for meta-analysis. In 15 comparisons, 486 patients were given intramuscular morphine, 10 mg, and 460 placebo *(Table 40)*. The size of the active treatment group in these trials varied between 9 patients and 51 patients (mean 33, median 30).

The placebo response rate (i.e. the proportion of patients given placebo experiencing at least 50% pain relief) varied from 0% to 47% (mean 15%), and the response rate for intramuscular morphine, 10 mg, was 7–93% (mean 46%; *Figure 37*). Of the 15 comparisons between intramuscular morphine, 10 mg, and placebo, eight showed it to be statistically superior to placebo and had a lower CI of the relative benefit greater than 1 *(Table 40)*. The pooled relative benefit was 2.8 (95% CI, 2.0–3.8).

The pooled NNT for intramuscular morphine, 10 mg, compared with placebo was 2.9 (95% CI, 2.6–3.6). Omitting a trial which included acute non-surgical pain[18] did not affect this result. The pooled NNT without this study was 3.1 (95% CI, 2.7–3.8). The NNT for trials in which fewer than the median number of patients were given morphine (that is, fewer than 32 patients treated) was 2.9 (95% CI, 2.3–4.1), the same as for larger trials (with 32 patients or more) – NNT of 3.0 (95% CI, 2.5–3.8).

Minor adverse effects occurred in 34% of patients given intramuscular morphine compared with 23% of patients given placebo. This was a significantly increased rate with a relative risk of 1.49 (95% CI, 1.09–2.04). Major adverse effects (drug-related study withdrawal) were rare (overall 1.2%) and did not differ between morphine and placebo *(Table 40)*.

Comment

Morphine is the archetypal analgesic for use in moderate or severe pain. It is also the 'gold standard' against which other injected analgesics are tested. It was surprising, therefore, that rigorous searching revealed so few placebo-controlled trials

TABLE 39 *Injected morphine in postoperative pain: patients, methods, outcomes and results of included studies*

Study	Condition and number of patients	Design, study duration and follow-up	Outcome measures	Treatment groups	Analgesic outcome results (morphine vs. placebo)	Withdrawals and adverse effects	Adverse events	Comment	Quality score
Beaver & Feise, 1976[16]	General and gynaecological surgery n = 96 Age: not given	RCT, double-blind, single dose, parallel group. Assessments by single nurse observer, hourly assessments up to 6 hour. Moderate to severe baseline pain.	Standard 4-point PI Standard 5-point PR 9-item tension/ anxiety questionnaire	1. i.m. morphine, 8 mg, n = 24 2. Placebo, n = 24 3. i.m. hydroxyzine, 100 mg, n = 24 4. i.m. hydroxyzine + morphine, n = 24	Morphine superior to placebo PID (see Figure 1).	Withdrawals not reported.	Sedative adverse effects: morphine 19/24, placebo 3/24.	Morphine better than placebo.	3
Brown, *et al.*, 1984[20]	Various surgical procedures n = 90 Age range: 18–68 years	RCT, double-blind, single dose, parallel group, double-dummy. Assessed by single nurse observer at 0, 0.5, 1 hour then hourly intervals for 6 hour. Medication taken when baseline pain was at least moderate.	PI (5-point) none, slight, moderate, severe, very severe PR (5-point) none, poor, fair, good, very good 50% PR at 6 hours Time to next analgesic	1. i.m. morphine, 10 mg, n = 30. 2. Placebo, n = 30 3. p.o. naproxen, 550 mg, n = 30	Significant difference between morphine and placebo for most outcomes. 1. % patients with > 50% PR: placebo 37%, morphine 77%, $p = 0.002$. 2. Mean hours in study (SD): placebo 3.2 (1.5), morphine 4.7 (1.4) 3. 6-hour mean SPID: placebo 2.1, morphine 5.5 4. 6-hour mean TOTPAR values not given. NB: 5-point SPID.	Patients remaining in study at 6 hours: placebo 5 (17%), morphine 13 (43%), $p = 0.01$. Early termination due to inadequate relief: morphine 16 (20%), placebo 25 (83%), $p = 0.01$.	No study withdrawals reported. NSD between treatments. Patients reporting one or more: placebo 14 (47%), morphine 22 (73%), naproxen 13 (43%).	NSAID better than morphine.	4
Brown, *et al.*, 1991[27]	Various surgical procedures n = 150 Age range: 18–66 years	Study 3 only. RCT, double-blind, single dose, double-dummy, parallel group. Assessed by single nurse observer at 0, 0.5, 1 hour then hourly intervals for 6 hours. Medication taken when baseline pain was at least moderate.	PI (5-point) 1 = none, 2 = slight, 3 = moderate, 4 = severe, 5 = very severe PR (5-point) 5 = none, 4 = poor, 3 = fair, 2 = good, 1 = very good Pain half gone at end of study, or at time of withdrawal	1. i.m. morphine 10 mg, n = 30 2. Placebo, n = 30 3. p.o. anirolac, 600 mg, n = 30 4. p.o. anirolac, 125 mg, n = 30 5. p.o. anirolac, 5 mg, n = 30	1. 6-hour mean SPID (SEM): morphine 5.2 (0.6), placebo 0.9 (1.0), $p < 0.003$. 2. 6 hour TOTPAR mean (SEM): morphine 17.7 (1.0), placebo 24.1 (1.1), $p < 0.003$. NB: wrong calculation, 50% TOTPAR not reported.	Completing 6-hour study: morphine 43%, placebo 30%.	No study withdrawals reported. NSD in % patients reporting: morphine 41%, placebo 23%.	NSAID better than morphine. Wrong calculations used for Figure 2 TOTPAR.	4
Campos & Solis, 1980[18]	Various acute and medical patients (2 part studies) n = 120 (each study) Age: adult	RCT, single dose, parallel group, double-blind. Assessments by single nurse observer at 0, 30 minutes, 1, 2, 3, 4, 5, 6 hours. Baseline pain at least moderate.	Standard 4-point PI 50% PR	1. i.m. morphine, 10 mg, n = 30 2. Placebo, n = 30 3. i.m. nefopam, 20 mg, n = 30 4. i.m. diphenhydramine, 20 mg, n = 30	1. SPID: morphine 12.32, placebo 6.48. 2. 50% PR: morphine 76%, placebo 49.6%.	Patients completing: morphine 28/30, placebo 23/40. Reporting no relief at 2 hours: morphine 1/30, placebo 6/30. Dropped out for other reasons: morphine 1, placebo 1.	Patients reporting: morphine 66, placebo 20.	Morphine better than placebo.	3
Davie, *et al.*, 1982[13]	Various day-surgery procedures n = 90 Age range: 23–69 years	RCT, single dose, parallel group, double-blind. Assessments by single observer at 0, 30 minutes, 1, 2, 3, 4, 5, 6 hours. Baseline pain at least moderate.	VAS PI 10 cm	1. i.v. morphine, 8 mg, n = 30 2. Placebo, n = 30 3. p.o. fenoprofen, 200 mg, n = 30	Morphine gave significantly less pain than placebo at all assessment times. No VAS SPID, but data can be calculated from table in text.	Withdrawals at 2 hours: morphine 3/30, placebo 20/30, fenoprofen 8/30.	Patients reporting: morphine 4, placebo 3, fenoprofen 2.	Morphine better than placebo.	4
de Andrade, *et al.*, 1994[14]	Orthopaedic surgery (hip and knee replacement) n = 176 Age: adult	RCT, double-blind, single dose, parallel group, double-dummy? Assessed by patients at 0, 0.5, 1 hour then hourly intervals for 6 hours. Medication taken when baseline pain was at least moderate.	PI (5-point) 0 = none, 1 = mild, 2 = moderate, 3 = severe, 4 = very severe VAS PI 100 mm 0–99 PR (5-point) 0 = none, 1 = a little, 2 = some, 3 = a lot, 4 = complete Patient and investigator global evaluation end of study	1. i.m. morphine, 10 mg, n = 51 2. i.m. morphine, 5 mg, n = 50 3. Placebo, n = 25 4. p.o. ketorolac, 10 mg, n = 50	1. 6-hour mean SPID (SD): morphine,10 mg, 7.2 (3.9); morphine, 5 mg, 6.7 (4.6); placebo 2.4 (3.8); $p < 0.01$. 2. 6-hour mean TOTPAR (SD): morphine, 10 mg, 14.0 (6.0); morphine, 5 mg, 12.4 (6.7); placebo 5.0 (7.3); $p < 0.01$.	105 remaining in study at 6 hours; morphine, 10 mg, 33; morphine, 5 mg, 32; placebo 7; $p < 0.01$.	5 patients withdrew: morphine 2; placebo 1. Morphine, 10 mg, significantly more than placebo, $p = 0.01$ (Table 4). Patients reporting: morphine, 10 mg, 26/51; morphine, 5 mg, 17/50; placebo 5/25.	NSAID comparable to morphine. Difference in sample size of active vs. placebo controls.	4

continued

TABLE 39 contd *Injected morphine in postoperative pain: patients, methods, outcomes and results of included studies*

Study	Condition and number of patients	Design, study duration and follow-up	Outcome measures	Treatment groups	Analgesic outcome results (morphine vs. placebo)	Withdrawals and adverse effects	Adverse events	Comment	Quality score
de Lia, *et al.*, 1986[24]	Gynaecological surgery n = 92 Age: not stated	RCT, double-blind, single dose, parallel group, double-dummy. Assessed by patients at 0, 1 hour then hourly intervals for 6 hours. Medication taken when baseline pain was at least moderate	Standard 4-point PI Investigator rating end of treatment: poor/fair = no effect, good/ excellent = effective PR derived score half gone Need for additional analgesia	1. Morphine, 10 mg, n = 30 2. Placebo, n = 30 3. p.o. flurbiprofen, 50 mg, n = 30	1. 6-hour mean SPID: morphine 5.17, placebo 1.80, $p < 0.01$. 2. Pain more than half gone at 6 hours: morphine 6 (20%), placebo 0 (0%). 3. 6-hour mean TOTPAR: morphine 16.98, placebo 10.80, $p < 0.01$.	Cumulative drop-out rate shown in Table II. Patients dropped-out at 6 hours: morphine 21 (70%), placebo 66 (74%).	Study withdrawals not stated. Patients reporting adverse effects: morphine 4, placebo 4.	NSAID comparable to morphine.	4
Fragen, et al., 1983[19]	Orthopaedic and major gynaecological surgery n = 139 Age range: 18–65 years	RCT, double-blind, single dose, parallel group, double-dummy. Assessed by single observer at 0, 15, 30 minutes, 1 hour then hourly intervals for 6 hours. Medication taken when baseline pain was at least moderate. Injections into deltoid muscle.	Standard 4-point PI PR (5-point): 1 = worse, 0 = none, 1 = a little, 2 = moderate, 3 = a lot, 4 = complete VAS PI 10 cm Patient rating end of treatment (4-point): 1 = poor, 2 = fair, 3 = good, 4 = excellent 50% PR (derived)	1. i.m. morphine, 10 mg, n = 36 2. Placebo, n = 35 3. i.m. ciramadol, 30 mg, n = 34 4. i.m. ciramadol, 60 mg, n = 34	1. 6-hour mean SPID: morphine 5.0, placebo 1.6. 2. 6-hour mean TOTPAR: morphine 10.1, placebo 3.63. Figure 1 shows % patients with > 50% PR at each time point.	Cumulative drop-out rate shown in Figure 2.	Study withdrawals not reported. Number of patients reporting not given. Nausea: morphine 5, placebo 1. Vomiting: morphine 4, placebo 1.	Morphine better than synthetic mixed agonist antagonist control.	3
Gravenstein, *et al.*, 1984[21]	Postoperative wound pain n = 160 Age range: 19–70 years	RCT, double-blind, single dose, parallel group. Assessed by more than 1 observer at 0, 15, 30 minutes, 1 hour then hourly intervals for 6 hours. Medication taken when baseline pain was at least moderate.	Standard 4-point PI PR (6-point): 1 = worse, 0 = none, 1 = slight, 2 = moderate, 3 = substantial, 4 = complete VAS PI 100 mm Patient rating of treatment (4-point): poor, fair, good, excellent Investigator rating of treatment (satisfactory, unsatisfactory)	1. i.m. morphine, 10 mg, n = 40 2. Placebo, n = 40 3. i.m. dezocine, 10 mg, n = 40 4. i.m. dezocine, 15 mg, n = 40	1. PI scores mean shown in Table 3. 2. PR mean scores shown in Figure 2. 3. Patients with moderate to complete PR by time shown in Table 3. Patients rating treatment as good or excellent: morphine 15/36, placebo 6/24.	Not given.	No study withdrawals reported. NSD reported between groups.	High dose of dezocine better than morphine.	3
Kaiko, *et al.*, 1987[17]	Acute postoperative pain n = 9/17 (completed crossover each treatment) Age range: 22–65 years	RCT, double-blind, single dose, crossover design. Assessed by nurse observer at 0, 30 mins, 1 hour then hourly intervals for 6 hours. Medication taken when baseline pain was at least moderate.	Standard 4-point PI PR (5-point) VAS PI VAS PR VAS mood	1. i.m. morphine, 10 mg, n = 9 2. Placebo, n = 9 3. Oral cocaine, 10 mg, n = 9 4. Oral cocaine, 10 mg, + morphine, n = 9	1. VAS TOTPAR: placebo 90, morphine 135, cocaine 70, cocaine + morphine 161. 2. VAS SPID: placebo 58, morphine 99, cocaine 35, cocaine + morphine 108.	9 of 17 post-operative patients completed crossover.	Patients reporting: placebo 2/12, morphine 6/13, cocaine 4/16. morphine + cocaine 7/13.	Morphine better than placebo.	3
Kantor, et al., 1981[30]	Postoperative and acute traumatic pain n = 250 Age range: 21–75 years	RCT, double-blind, 4 doses of same drug given over 2 days. Nurse observers, assessments at 0, 30 minutes, 1 hour then hourly intervals for 4 hours. Medication taken when baseline pain was at least moderate.	Standard 4-point PI Patients with ≥ 50% PR	1. i.m. morphine, 12.5 mg, n = 50 2. Placebo, n = 49 3. p.o. codeine, 90 mg, n = 50 4. p.o. pentazocine, 75 mg, n = 50 5. p.o. oxycodone compound (oxycodone hydrochloride, 4.5 mg, oxycodone terepthalate, 0.30 mg, aspirin, 224 mg, phenacetin, 160 mg, caffeine, 32 mg), n = 49	1. Derive SPID data for dose 1, day 1 from Figure 1. 2. SPID for day 1: morphine 4.90, placebo 1.20, pentazocine 3.43, codeine 3.63, oxycodone 4.35.	Not reported.	Total side-effects per dose: morphine (n = 14/50) 34; placebo (n = 4/49) 9; pentazocine (n = 12/50) ?; codeine (n = 11/50) 25; oxycodone (n = 2/49) 3.	Morphine better than placebo.	4

continued

TABLE 39 contd *Injected morphine in postoperative pain: patients, methods, outcomes and results of included studies*

Study	Condition and number of patients	Design, study duration and follow-up	Outcome measures	Treatment groups	Analgesic outcome results (morphine vs. placebo)	Withdrawals and adverse effects	Adverse events	Comment	Quality score
Lippmann, *et al.*, 1989[26]	Major abdominal and orthopaedic surgery n = 151 Age range: 18–65 years	RCT, double-blind, single dose, parallel group. Assessments at 0, 30 minutes, 1 hour then hourly intervals for 6 hours. Medication taken when baseline pain was at least moderate.	PI (4-point): 0 = none, 1 = slight, 2 = moderate, 3 = severe PR (scale not described) Additional analgesia Treatment effective, partially effective, ineffective (derived score)	1. i.m. morphine, 10 mg, n = 30 2. i.m. placebo, n = 30 3. i.m. tonazocine, 2 mg, n = 29 4. i.m. tonazocine, 4 mg, n = 30 5. i.m. tonazocine, 2 mg, n = 31	Morphine superior to placebo for most outcomes, as were other active treatments 1. SPID see Figure 4. 2. TOTPAR see Figure 3.	See Figure 6 for % remedicating by time.	Patients with none: morphine 10/30, placebo 25/30.	High dose of mixed agonist antagonist better than morphine; only extractable data from physician rating of effectiveness.	2
Morrison, *et al.*, 1986[25]	Major obstetric and gynaecological surgery n = 181 Age range: 19–65 years	RCT, double-blind, single dose, parallel group. Assessments at 0, 30 minutes, 1 hour then hourly intervals for 6 hours. Medication taken when baseline pain was at least moderate.	Standard 4-point PI PR: 1 = unchanged, worse, 2 ≤ half gone, 3 = half gone, 4 ≥ half gone, 5 = completely gone Investigator global: 1 = no effect, 2 = poor, 3 = fair, 4 = good, 5 = excellent Need for supplementary analgesia	1. i.m. morphine, 10 mg, n = 51 2. Placebo, n = 55 3. p.o. flurbiprofen, 50 mg, n = 53 4. p.o. zomepirac, 100 mg, n = 22	1. SPID mean 6-hour: morphine 10.7, placebo 5.07 (see Figure 2). 2. TOTPAR mean 6-hour: morphine 23.91, placebo 14.54 (see Figure 2) 3. Patients requesting additional analgesia: morphine 12/47, placebo 34/50.	Dropped-out at 6 hours: morphine 15, placebo 34.	1 patient reported adverse effect with morphine; no others.	No differences between NSAID and morphine. Zomepirac treatments incomplete as license withdrawn during study. Non-standard PR scale, no extractable data.	3
Nørholt, *et al.*, 1996[28]	Third molar extraction n = 252 Age range: 18–40 years	RCT, double-blind, single dose, double-dummy, parallel group. Assessments at 0, 15, 30, 45 minutes, 1 hour then hourly intervals for 8 hours. Medication taken when baseline pain was at least moderate.	PI (5-point): 0 = gone, 1 = slight, 2 = moderate, 3 = severe, 4 = unbearable PI (11-point) 0–10 PR (5-point): 0 = none to 4 = complete Patient rating of treatment 4 & 8 hours: 1–5 Onset of relief Duration of relief Remedication time	1. i.m. morphine, 10 mg, n = 37 2. i.m. morphine, 20 mg, n = 37 3. i.m. placebo, n = 37 4. i.m. lornoxicam, 4 mg, n = 33 5. i.m. lornoxicam, 8 mg, n = 38 6. i.m. lornoxicam, 16 mg, n = 38 7. i.m. lornoxicam, 20 mg, n = 37	1. 4-hour mean SPID (SD): morphine, 10 mg, 1.9 (2.7); morphine, 20 mg, 3.9 (3.0); placebo –5 (2.5). 2. 4-hour mean TOTPAR (SD): morphine, 10 mg, 5.1 (3.8); morphine, 20 mg, 8.8 (3.6); placebo 1.2 (2.3). 8-hour values available.	Median time to remedication, minutes (range): morphine, 10 mg, 185 (65–540); morphine, 20 mg, 540 (100–540); placebo 80 (30–540).	Patients reporting: morphine, 10 mg, 32/37; morphine, 20 mg, 37/37; placebo 22/37.	No differences between NSAID and morphine. Extractable data – global rating?	5
Pande, et al., 1996[29] 2 studies in this report; Study 2 only used.	Obstetric and gynaecological surgery n = 53 Age range: 26–61 years	RCT, double-blind, single dose, parallel group. Assessments at 0, 15, 30 minutes, 1 hour then hourly intervals for 6 hours. Medication taken when baseline pain was at least moderate.	PI 0–3 4 and 8 hour SPID PR 0–4 Pain reduced by half Patient rating of treatment at end of study	1. i.m. morphine, 10 mg, n = 14 2. i.m. placebo, n = 12 3. i.m. enadoline, 15 mg, n = 14 (kappa agonist) 4. i.m. enadoline, 25 mg, n = 13	1. 6-hour mean SPID values not given, only levels of significance. 2. 6-hour mean TOTPAR (SEM): morphine 2.7 (1.1), placebo 0.9 (1.2).	Study stopped early due to neuropsychiatric effects from enadoline. Patients completing study: morphine 1, placebo 0.	None reported for morphine or placebo.	No difference between enadoline and morphine. Problems with methods in study 2.	4
Pandit, *et al.*, 1985[22]	Orthopaedic, gynaecological and general surgery n = 190 Age range: 26–61 years	RCT, double-blind, single dose, parallel group. Single observer, assessments at 0, 15, 30 minutes, 1 hour then hourly intervals for 4 hours. Medication taken when baseline pain was at least moderate.	PI (3-point): mild, moderate, severe % PR: 0 = none, 1 ≤ 50%, 2 = 50%, 3 ≥ 50%, 4 = 100% Pain reduced by half	1. i.m. morphine, 10 mg, n = 39 2. i.m. placebo, n = 38 3. i.m. dezocine, 5 mg, n = 38 4. i.m. dezocine, 10 mg, n = 37 5. i.m. dezocine, 15 mg, n = 38	Morphine superior to placebo for some but not all outcomes. 1. 4-hour mean TOTPAR: morphine 5.8, placebo 3.1. 2. 50% relief at 4 hours: morphine 36.8%, placebo 11.4%. 3. Difference in proportion of patients with adequate relief at 2 and 4 hours only given for morphine vs. placebo.	Not stated.	Patients reporting: morphine 10%, placebo 8%.	Dezocine (mixed agonist antagonist) better than morphine on some outcomes.	4

continued

TABLE 39 contd *Injected morphine in postoperative pain: patients, methods, outcomes and results of included studies*

Study	Condition and number of patients	Design, study duration and follow-up	Outcome measures	Treatment groups	Analgesic outcome results (morphine vs. placebo)	Withdrawals and adverse effects	Adverse events	Comment	Quality score
Powell, 1985[23]	Orthopaedic, gynaecological and general surgery n = 160 Age range: 18–65 years	RCT, double-blind, multiple dose, parallel group, more than one observer, assessments at 0, 15, 30 minutes, 1 hour then hourly intervals for 6 hours. Medication taken when baseline pain was at least moderate.	Standard 4-point PI PR: –1 = worse, 0 = none, 1 = a little, 2 = moderate, 3 = a lot, 4 = complete Patient rating end of treatment: 1 = poor, 2 = fair, 3 = good, 4 = excellent	1. i.m. morphine, 10 mg, n = 40 2. i.m. placebo, n = 40 3. i.m. ciramadol, 30 mg, n = 40 4. i.m. ciramadol, 60 mg, n = 40	Morphine superior to placebo for all outcomes, as were other active treatments. 1. TOTPAR, see figures in text. 2. SPID, see figures. 3. VAS SPID, see figures. 4. Patient rating of good or excellent: morphine 48.7%, placebo 26%. Global: morphine 85%, placebo 26% (see text).	% remedicated at 6 hours: morphine 48.7, placebo 90.	Patients reporting: morphine 6 (15%), placebo 7 (18%).	Little difference between ciramadol and morphine. Extractable data: TOTPAR from figures in text.	4
van den Abeele & Camu, 1985[15]	Orthopaedic, gynaecological and general surgery n = 100 Age range: 18–65 years	RCT, double-blind, single, parallel group, medical observer, assessments at 0, 15, 30 minutes, 1 hour then hourly intervals for 6 hours. Medication taken when baseline pain was at least moderate.	Standard 4-point PI PR: –1 = worse, 0 = none, 1 = a little, 2 = moderate, 3 = a lot, 4 = complete VAS PI 10 cm Sedation: 0–3 Patient and investigator rating of treatment: poor, fair, good, excellent	1. i.m. morphine, 5 mg, n = 20 2. i.m. morphine, 10 mg, n = 20 3. i.m. placebo, n = 20 4. i.m. ciramadol, 30 mg, n = 40 5. i.m. ciramadol, 60 mg, n = 40	Morphine superior to placebo for most outcomes. 1. 6-hour TOTPAR: morphine, 10 mg, 13.6; morphine, 5 mg, 9.5; placebo 1.1. 2. 6-hour SPID: morphine, 10 mg, 9.3; morphine, 5 mg, 6.3; placebo 1.3. 3. 6-hour VAS SPID: morphine, 10 mg, 262.4; morphine, 5 mg, 132.0; placebo 37.9. Other data available.	Mean drop-out rate, see text.	Incidence: morphine, 10 mg, 0/2; morphine, 5 mg, 4/20; placebo 4/20.	Ciramadol better than morphine on some outcomes.	2

TABLE 40 *Analgesia and adverse effects of intramuscular morphine, 10 mg*

Trial (date order)	At least 50% PR with morphine	At least 50% PR with placebo	RB or RR (95% CI)	NNT (95% CI)
Campos, *et al.,* 1980[18]	28/30	14/30	2.0 (1.4–3.0)	2.1 (1.5–3.7)
van den Abeele & Camu, 1983[15]	15/20	2/20	7.5 (2.0–28.6)	1.5 (1.1–2.4)
Fragen, *et al.,* 1983[19]	17/36	4/35	4.1 (1.5–11.1)	2.8 (1.8–6.1)
Brown, *et al.,* 1984[20]	23/30	11/29	2.0 (1.2–3.4)	2.6 (1.6–6.5)
Gravenstein, 1984[21]	10/40	0/40	101 (0.2–> 250)	4.0 (2.6–8.8)
Pandit, *et al.,* 1985[22]	8/39	2/38	3.9 (0.9–17.2)	6.7 (3.4–138)
Powell, 1985[23]	11/39	0/40	114 (0.2–> 250)	3.6 (2.4–7.1)
de Lia, *et al.,* 1986[24]	15/30	5/30	3.0 (1.3–7.2)	3.0 (1.8–9.1)
Morrison, *et al.,* 1986[25]	47/51	25/55	2.0 (1.5–2.7)	2.1 (1.6–3.2)
Kaiko, *et al.,* 1987[17]	2/9	1/9	2.0 (0.2–18.8)	9.1 (2.2–∞)
Lippmann, *et al.,* 1989[26]	4/30	0/30	41 (0.1–> 250)	7.7 (3.9–85)
Brown, *et al.,* 1991[27]	17/30	6/30	2.8 (1.3–6.2)	2.7 (1.7–7.2)
de Andrade, *et al.,* 1994[14]	34/51	4/25	4.2 (1.7–10.5)	2.0 (1.4–3.2)
Nørholt, *et al.,* 1996[28]	9/37	0/37	91 (0.2–> 250)	4.2 (2.6–9.5)
Pande, *et al.,* 1996[29]	1/14	0/12	9.4 (0.0–> 250)	14.3 (4.9–∞)
Combined analgesic data	241/486	74/460	2.8 (2.0–3.8)	2.9 (2.6–3.6)
Trials with < 32 treated patients	101/293	40/198	2.2 (1.8–2.8)	2.9 (2.3–4.1)
Trials with > 32 treated patients	136/293	35/270	4.0 (1.6–9.8)	3.0 (2.5–3.8)
Minor adverse effects	108/320	68/295	1.49 (1.09–2.04)	9.1 (5.6–27.7)
Major adverse effects	2/334	6/304	0.31 (0.07–1.38)	

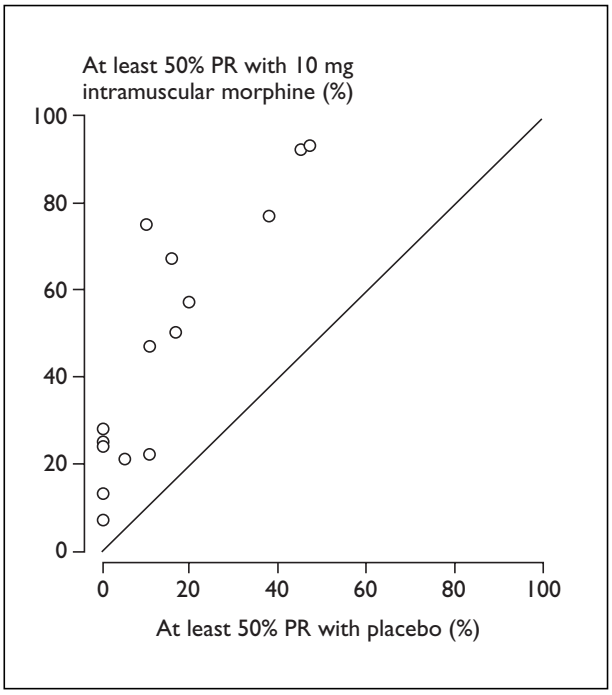


FIGURE 37 *Relationship of the proportion of patients achieving at least 50% pain relief with intramuscular morphine, 10 mg, to the proportion obtaining at least 50% pain relief with placebo in 15 trials*

in which morphine was given by intravenous, intramuscular or subcutaneous injection, and in which single-dose analgesic efficacy was tested using standard, validated methods. No subcutaneous studies were found and only one intravenous study, and only for intramuscular morphine, 10 mg, was there sufficient information (494 treated patients) for it to be pooled for meta-analysis. No studies of diamorphine were found which met the criteria.

A single intramuscular dose of morphine, 10 mg, had an NNT of 2.9 for at least 50% pain relief compared with placebo. This means that one in every three patients with pain of moderate to severe intensity will experience at least 50% pain relief with morphine which they would not have had with placebo. Sensitivity analysis found that size of trial did not make a difference *(Table 40)*. Sensitivity analysis was not performed for quality of trials, since all but two reports had quality scores of 3 or more. Overestimation of the effect of treatment has been shown in trials with quality scores of 2 or less using the same validated quality scale as here.[40]

The NNT for morphine can be compared with those of other analgesics from similar meta-analyses in which the efficacy of analgesics was compared with placebo in patients with moderate or severe postoperative pain. While there is as yet no comparable information available for other injected analgesics, the NNT of 2.9 (95% CI, 2.6–3.8) for intramuscular morphine, 10 mg, can be compared with those obtained for oral tramadol, 100 mg, (4.8 (95% CI, 3.4–8.2)),[41] for oral paracetamol, 1000 mg, (4.6 (95% CI, 3.9–5.4)), for paracetamol, 600/650 mg, plus codeine, 60 mg (3.1 (95% CI, 2.6–3.8)),[42] and for ibuprofen, 400 mg, (2.7 (95% CI, 2.5–3.0)). The equivalence of the NNTs for oral NSAIDs and intramuscular morphine, 10 mg, is supported by the repeated failure to separate them in analgesic trials.[43,44] A crucial issue here is dose. Clearly with opioids, there should be dose titration against effect. The NNT value of 2.9 is for 10 mg of intramuscular morphine; giving 20 mg improved the NNT value.[28]

Rank ordering of analgesics in this way is potentially less accurate than taking the relative efficacy of the individual drugs from within one very large trial with a single randomisation. In the absence of such head-to-head' comparisons, the authors consider that this indirect ranking, the relative efficacy of the drugs against placebo, is helpful in making clinical decisions. The trials used to produce NNTs for analgesics compared with placebo are all single-dose, postoperative, randomised and double-blind. The patients must have had moderate to severe pain before being treated and standard measures of pain were required. These uniform quality standards and patient selection criteria allow a credible indirect ranking of efficacy to be made. Internal validity is demonstrated in the ranking by the dose–response relationships obtained for analgesics, with better analgesia (lower NNT) obtained with higher doses *(Figure 56)*. External validity will come when there are direct (head-to-head) comparisons which confirm the rank order in the indirect table.

At first sight the fact that the analgesia from intramuscular morphine, 10 mg, is no better than the analgesia from a therapeutic dose of oral NSAID is surprising.[45] Injected drugs are generally thought of as more 'powerful' than oral drugs. In reality, there is a considerable body of direct evidence that confirms the indirect ranking. For many years investigators have been unable to distinguish the analgesia resulting from intramuscular morphine, 10 mg, and oral NSAID (where comparisons were within the same trial and, hence, randomised).

This is a clinically useful observation for patients who can swallow and who have no contraindication to NSAID. Oral NSAID appears to be the best analgesic choice. There is no advantage to giving that dose of NSAID by a suppository or injection.[46] If

the patient can swallow but speedy analgesia is required, then intravenous rather than intramuscular analgesia seems more logical. If the patient cannot swallow, then intramuscular morphine, 10 mg, gives analgesia equivalent to oral NSAID, and doubling the dose does indeed increase the analgesia.[28] There is not, as yet, a ranking of injected NSAID compared with injected opioid.

References

1. McQuay HJ, Carroll D, Moore RA. Injected morphine in postoperative pain; a quantitative systematic review. *J Pain Symptom Manage*. In press.
2. Jadad AR, Carroll D, Moore A, McQuay H. Developing a database of published reports of randomised clinical trials in pain research. *Pain* 1996;**66**:239–46.
3. Jadad AR, Moore RA, Carroll D, Jenkinson C, Reynolds DJM, Gavaghan DJ, *et al.* Assessing the quality of reports of randomized clinical trials: is blinding necessary? *Control Clin Trials* 1996;**17**:1–12.
4. Cooper SA. Single-dose analgesic studies: the upside and downside of assay sensitivity. *Adv Pain Res Ther* 1991;**18**:117–24.
5. Moore A, McQuay H, Gavaghan D. Deriving dichotomous outcome measures from continuous data in randomised controlled trials of analgesics. *Pain* 1996;**66**:229–37.
6. Moore A, Moore O, McQuay H, Gavaghan D. Deriving dichotomous outcome measures from continuous data in randomised controlled trials of analgesics: use of pain intensity and visual analogue scales. *Pain* 1997;**69**:311–15.
7. Moore A, McQuay H, Gavaghan D. Deriving dichotomous outcome measures from continuous data in randomised controlled trials of analgesics: verification from independent data. *Pain* 1997;**69**:127–30.
8. DerSimonian R, Laird N. Meta-analysis of clinical trials. *Control Clin Trials* 1986;**7**:177–88.
9. Gardner MJ, Altman DG. Confidence intervals rather than p values: estimation rather than hypothesis testing. *BMJ* 1986;**292**:746–50.
10. L'Abbé KA, Detsky AS, O'Rourke K. Meta-analysis in clinical research. *Ann Intern Med* 1987;**107**:224–33.
11. Laupacis A, Sackett DL, Roberts RS. An assessment of clinically useful measures of the consequences of treatment. *N Engl J Med* 1988;**318**:1728–33.
12. Cook RJ, Sackett DL. The number needed to treat: a clinically useful measure of treatment effect. *BMJ* 1995;**310**:452–4.
13. Davie IT, Slawson KB, Burt RA. A double blind comparison of parenteral morphine, placebo, and oral fenoprofen in management of postoperative pain. *Anesth Analg* 1982;**61**:1002–5.
14. de Andrade JR, Maslanka MA, Maneatis T, Bynum L, Burchmore M. The use of ketorolac in the management of postoperative pain. *Orthopedics* 1994;**17**:157–66.
15. van den Abeele G, Camu F. Comparative evaluation of ciramadol (WY 15.705), morphine and placebo for treatment of postoperative pain. *Acta Anaesthesiol Belg* 1985;**36**:97–110.
16. Beaver WT, Feise G. Comparison of the analgesic effects of morphine, hydroxyzine, and their combination in patients with postoperative pain. *Adv Pain Res Ther* 1976;**1**:553–7.
17. Kaiko R, Kanner R, Foley K, Wallenstein S, Canel A, Rogers A, *et al.* Cocaine and morphine interaction in acute and chronic cancer pain. *Pain* 1987;**31**:35–45.
18. Campos VM, Solis EL. The analgesic and hypothermic effects of nefopam, morphine, aspirin, diphenhydramine, and placebo. *J Clin Pharmacol* 1980;**20**:42–9.
19. Fragen RJ, Kouzmanoff C, Caldwell NJ. Intramuscularly administered ciramadol for management of postoperative pain: a comparative study. *J Clin Pharmacol* 1983;**23**:219–26.
20. Brown CR, Sevelius H, Wild V. A comparison of single doses of naproxen sodium, morphine sulfate, and placebo in patients with postoperative pain. *Curr Ther Res* 1984;**35**:511–18.
21. Gravenstein JS. Dezocine for postoperative wound pain. *Int J Clin Pharmacol Ther Toxicol* 1984;**22**:502–5.
22. Pandit SK, Kothary SP, Pandit UA, Kunz NR. Double blind placebo controlled comparison of dezocine and morphine for post operative pain relief. *Can Anaesth Soc J* 1985;**32**:583–91.
23. Powell W. A double blind comparison of multiple intramuscular doses of ciramadol, morphine, and placebo for the treatment of postoperative pain. *Anesth Analg* 1985;**64**:1101–7.
24. de Lia JE, Rodman KC, Jolles CJ. Comparative efficacy of oral flurbiprofen, intramuscular morphine sulfate, and placebo in the treatment of gynecologic postoperative pain. *Am J Med* 1986;**80**:60–4.
25. Morrison J, Harris J, Sherrill J, Heilman C, Bucovaz E, Wiser W. Comparative study of flurbiprofen and morphine for postsurgical gynecologic pain. *Am J Med* 1986;**80**:55–9.
26. Lippmann M, Mok M, Farinacci J, Lee J. Tonazocine mesylate in postoperative pain patients: a double blind placebo controlled analgesic study. *J Clin Pharmacol* 1989;**29**:373–8.
27. Brown CR, Schwartz KE, Wild VM, Koshiver JE. Comparison of anirolac with morphine and placebo for postoperative pain. *Curr Ther Res* 1991;**50**:379–85.

28. Norholt S, SindetPedersen S, Larsen U, Bang U, Ingerslev J, Nielsen O, *et al.* Pain control after dental surgery: a double-blind, randomised trial of lornoxicam versus morphine. *Pain* 1996;**67**:335–43.

29. Pande AC, Pyke RE, Greiner M, Wideman GL, Benjamin R, Pierce MW. Analgesic efficacy of enadoline versus placebo or morphine in postsurgical pain. *Clin Neuropharmacol* 1996;**19**:451–6.

30. Kantor TG, Hopper M, Laska E. Adverse effects of commonly ordered oral narcotics. *J Clin Pharmacol* 1981;**21**:1–8.

31. Rice A, Lloyd J, Miller C, Bullingham RE, O'Sullivan G. A double-blind study of the speed of onset of analgesia following intramuscular administration of ketorolac tromethamine in comparison to intramuscular morphine and placebo. *Anaesthesia* 1991;**46**:541–4.

32. Whitehead EM, O'Sullivan GM, Lloyd J, Bullingham RES. A new method for rate of analgesic onset: two doses of intravenous morphine compared with placebo. *Clin Pharmacol Ther* 1992;**52**:197–204.

33. Verborgh C, Camu F. Post surgical pain relief with zero order intravenous infusions of meptazinol and morphine: a double blind placebo controlled evaluation of their effects on ventilation. *Eur J Clin Pharmacol* 1990;**38**:437–42.

34. Maslanka MA, de Andrade JR, Maneatis T, Bynum L, DiGiorgio E. Comparison of oral ketorolac, intramuscular morphine, and placebo for treatment of pain after orthopedic surgery. *South Med J* 1994;**87**:506–13.

35. Brown CR, Moodie JE, Dickie G, Wild VM, Smith BA, Clarke PJ, *et al.* Analgesic efficacy and safety of single dose oral and intramuscular ketorolac tromethamine for postoperative pain. *Pharmacotherapy* 1990;**10**:59–70S.

36. Cohen RI, Edwards W T, Kezer EA, Ferrari DA, Liland AE, Smith ER. Serial intravenous doses of dezocine, morphine, and nalbuphine in the management of postoperative pain for outpatients. *Anesth Analg* 1993;**77**:533–9.

37. Lasagna L, Mosteller F, von Felsinger JM, Beecher HK. A study of the placebo response. *Am J Med* 1954:770–9.

38. Levine J, Gordon N, Smith R, Fields H. Analgesic responses to morphine and placebo in individuals with post-operative pain. *Pain* 1981;**14**:379–88.

39. Foley WL, Edwards RC, Jacobs L3. Patient-controlled analgesia: a comparison of dosing regimens for acute postsurgical pain. *J Oral Maxillofac Surg* 1994;**52**:155–9; discussion 159–60.

40. Khan KS, Daya S, Jadad AR. The importance of quality of primary studies in producing unbiased systematic reviews. *Arch Intern Med* 1996;**156**:661–6.

41. Moore RA, McQuay HJ. Single-patient data meta-analysis of 3453 postoperative patients: oral tramadol versus placebo, codeine and combination analgesics. *Pain* 1997;**69**:287–94.

42. Moore A, Collins S, Carroll D, McQuay H. Paracetamol with and without codeine in acute pain: a quantitative systematic review. *Pain* 1997;**70**:193–201.

43. Mansfield M, Firth F, Glynn C, Kinsella J. A comparison of ibuprofen arginine with morphine sulphate for pain relief after orthopaedic surgery. *Eur J Anaesthesiol* 1996;**13**:492–7.

44. Schachtel BP, Thoden WR, Baybutt RI. Ibuprofen and acetaminophen in the relief of postpartum episiotomy pain. *J Clin Pharmacol* 1989;**29**:550–3.

45. Collins SL, Moore A, McQuay HJ, Wiffen PJ. Oral ibuprofen and diclofenac in postoperative pain: a quantitative systematic review. *Eur J Pain*. In press.

46. McQuay HJ, Justins D, Moore RA. Treating acute pain in hospital. *BMJ* 1997;**314**:1531–5.

Chapter 13

Dihydrocodeine in postoperative pain

Summary

The aim of this review was to determine the analgesic efficacy and adverse effects of oral and injectable dihydrocodeine from single dose studies in moderate to severe postoperative pain. Published studies were identified by searching electronic databases and checking reference lists of retrieved reports. Summed pain relief and pain intensity data were extracted and converted to dichotomous information yielding the number of patients with at least 50% pain relief. This was used to calculate the relative benefit and NNT for one patient to achieve at least 50% pain relief.

In three reports (194 patients) oral dihydrocodeine was compared with placebo and in one (120 patients) dihydrocodeine, 30 mg or 60 mg, was compared with ibuprofen, 400 mg. For a single dose of dihydrocodeine, 30 mg, in moderate to severe postoperative pain, the NNT for at least 50% pain relief was 9.7 (95% CI, 4.5–∞) when compared with placebo over a period of 4–6 hours. Pooled data showed no significant difference in adverse effect incidence for dihydrocodeine, 30 mg, compared with placebo.

The 95% CIs of the NNT included no benefit of dihydrocodeine, 30 mg, over placebo. A statistical superiority for ibuprofen, 400 mg, over dihydrocodeine, 30 mg or 60 mg, was shown.

This chapter of the review has been published in full by Edwards and colleagues.[1]

Introduction

Opioids are extensively used in the management of pain and are believed capable of relieving severe pain more effectively than NSAIDs.[2] The aim of this quantitative systematic review was to assess the efficacy and safety of a single dose of oral dihydrocodeine in the management of postoperative pain of moderate to severe intensity.

Dihydrocodeine is a synthetic opioid analgesic developed in the early 1900s. Its structure and pharmacokinetics are similar to that of codeine[3] and it is used for the treatment of postoperative pain or as an antitussive. In 1995, nearly one-tenth of all analgesic prescriptions (opiate, non-opiate and NSAID) issued in the UK were for dihydrocodeine.[4] The proportion of dihydrocodeine used for the treatment of postoperative pain is not known.

Methods

A search was undertaken for RCTs of dihydrocodeine in postoperative pain which covered MEDLINE (1966–February 1997), EMBASE (1980–97), the Cochrane Library (January 1997), Biological Abstracts (1985–97), and the Oxford Pain Relief Database (1950–94).[5] The terms 'dihydrocodeine', 'random*', 'clinical trial', 'trial', analgesi*', 'pain' and 36 brand names and preparations[6] were used in a broad free-text search without restriction to language. Additional reports were identified from reference lists of retrieved articles. Unpublished data were not sought.

Included reports

The inclusion criteria used were:

- full journal publication
- postoperative pain
- postoperative administration
- adult patients
- baseline pain of moderate to severe intensity
- double-blind design
- random allocation to treatment groups which included dihydrocodeine and placebo.

Pain outcomes used were TOTPAR or SPID over 4–6 hours or sufficient data provided to allow their calculation. Pain measures allowed for the calculation of TOTPAR were a standard five-point pain relief scale (none, slight, moderate, good, complete), and for SPID a standard four-point pain intensity scale (none, mild, moderate, severe).

Data extraction and analysis

Extracted from each study were:

- the number of patients treated
- the mean TOTPAR or SPID
- study duration
- the dose of dihydrocodeine
- information on adverse effects.

Mean TOTPAR or SPID values were converted to % maxTOTPAR or % maxSPID by division into the calculated maximum value.[7] The referenced equations were used to estimate the proportion of patients achieving at least 50% maxTOTPAR.[8,9] This was then converted to the number of patients achieving at least 50% maxTOTPAR by multiplying by the total number of patients in the treatment group. The number of patients with at least 50% maxTOTPAR was then used to calculate estimates of relative benefit and NNT.

Estimates of relative benefit and risk, with 95% CIs, were calculated using a random effects model.[10] Homogeneity was assumed when $p > 0.1$. A statistically significant benefit of active treatment over control was assumed when the CI did not include 1. A statistically significant benefit of control over active treatment was assumed when the upper limit of the 95% CI of the relative benefit was < 1. NNT and NNH with 95% CIs were calculated.[11] The 95% CI of the NNT indicates no benefit of one treatment over the other when the upper limit includes infinity.

Results

A total of 48 published reports of dihydrocodeine in postoperative pain were identified, of which two could not be obtained from the British Library. Of the retrieved reports, 18 studies were not randomised and were excluded, leaving 28 randomised studies. Of these, two included other pain conditions, five had no extractable pain outcome data, seven were not double-blind, four used dihydrocodeine as a rescue analgesic only, and six did not specify baseline pain of moderate to severe intensity. These reports were also excluded. Details of the included studies are given in *Table 41*.

TABLE 41 *Dihydrocodeine in postoperative pain: summary of included trials*

Study	Condition and number of patients	Study design, duration and follow-up	Outcome measures	Dosing regimen	Analgesic outcome results	Remedication	Withdrawals and exclusions	Adverse effects
Dihydrocodeine versus placebo								
Frame, *et al.*, 1989[12]	Impacted third molar removal n = 148 Age: adult	RCT, double-blind, single oral dose, parallel groups. Assessed at 0.5, 1 hour and then hourly for 5 hours. Medication taken when pain of moderate to severe intensity.	PI (9-point scale), non-standard PR (5-point scale) standard	Dihydrocodeine, 30 mg (n = 49); placebo (n = 50)	Dihydrocodeine 30 mg not significantly different to placebo. 4-hour TOTPAR: dihydrocodeine, 30 mg, 0.5; placebo 0.3	Remedication allowed at 2 hours. If remedicated patients withdrawn and PR set to zero for all further time points.	18 withdrew: 9 insufficient pain, 7 did not return assessment forms, 1 did not complete assessment forms, 1 postoperative complications.	No serious adverse effects reported; no patients withdrew as result. Dihydrocodeine, 30 mg: 1/49 with 1 adverse effect. Placebo: 1/50 with 3 adverse events.
Galasko, et al., 1989[13]	Orthopaedic surgery n = 89 Age range: 18–80 years	RCT, double-blind, multiple oral dose, parallel groups. Assessed at 0.5, 1 hour and then hourly for 6 hours. Medication taken when pain of moderate to severe intensity.	PI (5-point scale), non-standard PR (5-point scale), standard VAS, 100 mm ('no pain' to 'worst pain I have ever felt')	Dihydrocodeine, 30 mg (n = 30); placebo (n = 28)	Dihydrocodeine not significantly different to placebo. Mean TOTPAR at 6 hours: dihydrocodeine 11.3, placebo 11.1	Multiple dose study; second dose given as required. If remedicated, patients excluded from analysis.	9 withdrew because of inadequate analgesia after first dose. Dihydrocodeine, 30 mg (n = 3); placebo (n = 6).	No patients experienced adverse effects in single dose analysis.
McQuay, *et al.*, 1985[14]	Minor day-case surgery (general) n = 54 Age: adult	RCT, double-blind, multiple oral dose, parallel groups. Assessed at 0.5, 1 hour and then hourly for 4 hours. Medication taken when pain of moderate to severe intensity.	PI (4-point scale), standard PR (5-point scale), standard VAS, 100 mm	Dihydrocodeine, 30 mg (n = 18); placebo (n = 19)	4-hour SPID and TOTPAR presented. TOTPAR: dihydrocodeine significantly better than placebo ($p < 0.05$); dihydrocodeine, 30 mg, 6.5; placebo 3.2	Allowed after 1 hour. If remedicated patients initial PI and PR scores used for all further time points.	Single dose analysis: all adverse effects mild, no patients withdrew as result. NSD between dihydrocodeine and placebo.	Dihydrocodeine, 30 mg: 6/18 with 6 adverse events. Placebo: 3/19 with 3 adverse events.
Dihydrocodeine versus ibuprofen								
McQuay, *et al.*, 1993[15]	Lower third molar removal n = 68 Age: adult	RCT, double-blind, multiple oral dose, cross-over design. Self-assessed at 0.5, 1 hour and then hourly for 6 hours. Medication allowed when pain of moderate to severe intensity.	PI (4-point scale), standard PR (5-point scale), standard Global rating (5-point scale), standard	Dihydrocodeine, 30 mg (n = 40); dihydrocodeine, 60 mg (n = 40); placebo (n = 40)	TOTPAR at 6 hours: dihydrocodeine, 30 mg, 3.3; dihydrocodeine, 60 mg, 4.7; ibuprofen, 400 mg, 10.0 Ibuprofen significantly better than dihydrocodeine, 30 mg or 60 mg, ($p < 0.01$)	If remedicated at 6 hours, initial PI score and PR score of zero used for all further time points.	3 patients withdrew.	Single-dose adverse effects data were not presented.

Four studies met the inclusion criteria: three were placebo-controlled and one used ibuprofen, 400 mg, as an active control. All four studies examined the effects of oral dihydrocodeine. Three trials[12–14] compared dihydrocodeine, 30 mg, with placebo and one[15] compared dihydrocodeine, 30 mg or 60 mg, with ibuprofen, 400 mg.

Oral dihydrocodeine versus placebo

No reports comparing dihydrocodeine, 60 mg, with placebo met our inclusion criteria. Three reports compared dihydrocodeine tartrate, 30 mg, (91 patients) with placebo (85 patients). One trial investigated dental pain,[12] one orthopaedic pain,[13] and one pain following minor day-case surgery.[14]

The proportion of patients experiencing at least 50% pain relief with dihydrocodeine varied between 14% and 50%, with a mean value of 35%. The proportion of patients experiencing at least 50% pain relief with placebo varied between 5% and 50%, with a mean of 23% *(Figure 38)*. The data sets were homogeneous ($p = 0.12$). Dihydrocodeine, 30 mg, was not significantly different from placebo, relative benefit 1.7 (95% CI, 0.7–4.0) *(Table 42)*. For a single dose of dihydrocodeine, 30 mg, compared with placebo the NNT was 9.7 (95% CI, 4.5–∞) for at least 50% pain relief over a period of 4–6 hours in postoperative pain of moderate to severe intensity.

Adverse effects

Details of adverse effects are given in *Table 43*. The incidence of adverse effects with dihydrocodeine was not significantly different from placebo. All adverse effects were mild and transient in nature and no patients withdrew as a result.

Oral dihydrocodeine vs ibuprofen

In one study,[15] the efficacy and safety of either dihydrocodeine tartrate, 30 mg (40 patients) or 60 mg (40 patients), was compared with ibuprofen, 400 mg (40 patients), in dental pain.

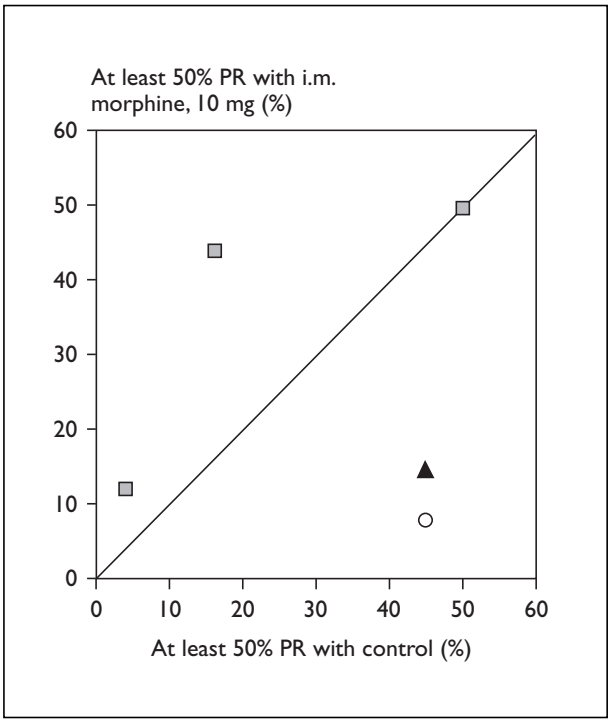


FIGURE 38 *Trials of oral dihydrocodeine in postoperative pain (■, dihydrocodeine, 30 mg, vs. placebo; ○, dihydrocodeine, 30 mg, vs. ibuprofen, 400 mg; ▲, dihydrocodeine, 60 mg, vs. ibuprofen, 400 mg)*

The proportion of patients experiencing at least 50% pain relief with dihydrocodeine, 30 mg, was 8%, with dihydrocodeine, 60 mg, it was 15% and with ibuprofen, 400 mg, (active control) it was 45% *(Figure 38)*. A statistical superiority of ibuprofen, 400 mg, over dihydrocodeine, 30 mg, and dihydrocodeine, 60 mg, was shown, with relative benefit values of 0.2 (95% CI, 0.1–0.5) and 0.3 (95% CI, 0.2–0.8), respectively.

Ibuprofen, 400 mg, was significantly better than dihydrocodeine, 30 mg, or dihydrocodeine, 60 mg,

TABLE 42 *Summary of relative and NNT for trials of dihydrocodeine against placebo and ibuprofen, 400 mg*

Number of trials	Dose of dihydrocodeine	Number of patients with > 50% PR: dihydrocodeine	Number of patients with > 50% PR: placebo or ibuprofen, 400 mg	RB (95% CI)	NNT (95% CI)
Versus placebo					
3	30 mg	29/97	19/97	1.7 (0.7, 4.0)	9.7 (4.5, ∞)
Versus ibuprofen, 400 mg					
1	30 mg	3/40	18/40	0.2 (0.1, 0.5)	–2.7 (–1.8, –5)
1	60 mg	6/40	18/40	0.3 (0.2, 0.8)	–3.3 (–2.1, –9)

Negative NNTs in the comparison with ibuprofen mean that ibuprofen is better than dihydrocodeine.

(Figure 38). When compared with ibuprofen, 400 mg, the NNT for a single dose of dihydrocodeine, 30 mg, was –2.7 (95% CI, –1.8, –5) for at least 50% pain relief over a period of 4–6 hours in postoperative pain of moderate to severe intensity *(Table 42)*. Similarly, for a single dose of dihydrocodeine, 60 mg, the NNT was –3.3 (95% CI, –2.1, –9) for at least 50% pain relief over a period of 4–6 hours.

Adverse effects

No single dose adverse effect data were presented.[15]

Comment

Dihydrocodeine is the second most commonly prescribed opioid in England, with 1.5 million prescriptions issued for dihydrocodeine tartrate tablets alone in 1995. This increased to 1.6 million in 1996. No papers were found which investigated injected dihydrocodeine in the evaluation of postoperative pain with standard analgesic measurement methods.

For a single dose of oral dihydrocodeine tartrate, 30 mg, compared with placebo the NNT was 9.7 (95% CI, 4.5–∞) for at least 50% pain relief over a period of 4–6 hours in postoperative pain of moderate to severe intensity. This means that one in every ten patients with moderate to severe postoperative pain would experience at least 50% pain relief with dihydrocodeine, 30 mg, who would not have done so with placebo. However, the estimate of relative benefit showed no significant difference between dihydrocodeine, 30 mg, and placebo.

A rank order of single dose analgesic efficacy in postoperative pain of moderate to severe intensity has been established by comparing orally administered analgesics from methodologically similar studies in other chapters in this review. A number of analgesics demonstrated greater efficacy than for dihydrocodeine, 30 mg, although the NNT 95% CIs for many of these overlap. The 95% CIs for ibuprofen, 200 mg (2.8–4.0) and 400 mg (2.5–3.0), and diclofenac, 50 mg (2.0–2.7) do not overlap with those those for dihydrocodeine, 30 mg, indicating greater analgesic efficacy.

This rank order of relative efficacy against placebo is supported by a head-to-head comparison with ibuprofen. The analgesic efficacy of a single dose of oral dihydrocodeine, 30 mg or 60 mg, was significantly inferior to ibuprofen, 400 mg. For a single dose of dihydrocodeine, 30 mg, compared with ibuprofen, 400 mg, the NNT was –2.7 (95% CI, –1.8, –5) for at least 50% pain relief over a period of 4–6 hours in postoperative pain of moderate to severe intensity. This means that for every three patients with moderate to severe postoperative pain treated with ibuprofen, 400 mg, one will experience at least 50% pain relief who would not have done if given dihydrocodeine, 30 mg.

Similarly, for a single dose of dihydrocodeine, 60 mg, compared with ibuprofen, 400 mg, the NNT was –3.3 (95% CI, –2.1, –9) over a period of 4–6 hours. So, one in every three patients with moderate to severe postoperative pain treated with ibuprofen, 400 mg, would experience at least 50% pain relief who would not have done if given dihydrocodeine, 60 mg.

Nausea, vomiting, headache, dizziness, drowsiness and confusion were the most commonly reported adverse effects for a single dose of oral dihydrocodeine, 30 mg, when compared with placebo. The incidence of adverse effects was not significantly different for dihydrocodeine, 30 mg, than for placebo *(Table 43)*.

Our results suggest dihydrocodeine to be less effective than other analgesics when administered as a single oral dose. Few of the retrieved reports investigating oral dihydrocodeine met the criteria

TABLE 43 *Summary of adverse effects of oral dihydrocodeine and placebo*

Number of trials	Adverse effect	Number of patients with adverse effects: dihydrocodeine	Number of patients with adverse effects: placebo	RR (95% CI)	NNH (95% CI)
3	Nausea or vomiting	7/97	0/97	25 (0.7–907)	N/A
3	Headache	3/97	0/97	1.05 (0.3–4.4)	N/A
3	Dizziness, drowsiness or confusion	5/97	1/97	4.2 (0.6–28)	N/A

N/A: Not calculated because NSD from placebo was shown for RR.

for inclusion in this quantitative systematic review. This resulted in very little patient data being available for analysis, particularly for dihydrocodeine, 60 mg, which is often the preferred dose. Administering dihydrocodeine in multiple doses may improve its analgesic efficacy but may also increase the incidence of adverse effects.

References

1. Edwards JE, McQuay HJ, Moore RA. Systematic review: dihydrocodeine in postoperative pain. Cochrane Library. In press.
2. Alexander JI, Hill RG. Postoperative pain. London: Blackwell Scientific Publications, 1987.
3. Rowell FJ, Seymour RA, Rawlins MD. Pharmacokinetics of intravenous and oral dihydrocodeine and its metabolites. *Eur J Clin Pharmacol* 1983;**25**:419–24.
4. Government Statistical Service. Prescription cost analysis for England 1995. London: Department of Health, 1996.
5. Jadad AR, Carroll D, Moore A, McQuay H. Developing a database of published reports of randomised clinical trials in pain research. *Pain* 1996;**66**:239–46.
6. Reynolds JEF. Martindale: the extra pharmacopoeia. 30th edition. London: Pharmaceutical Press, 1993:292–314.
7. Cooper SA. Single-dose analgesic studies: the upside and downside of assay sensitivity. *Adv Pain Res Ther* 1991;**18**:117–24.
8. Moore A, McQuay H, Gavaghan D. Deriving dichotomous outcome measures from continuous data in randomised controlled trials of analgesics: verification from independent data. *Pain* 1997;**69**:127–30.
9. Moore A, Moore O, McQuay H, Gavaghan D. Deriving dichotomous outcome measures from continuous data in randomised controlled trials of analgesics: use of pain intensity and visual analogue scales. *Pain* 1997;**69**:311–15.
10. DerSimonian R, Laird N. Meta-analysis of clinical trials. *Control Clin Trials* 1986;**7**:177–88.
11. Cook RJ, Sackett DL. The number needed to treat: a clinically useful measure of treatment effect. *BMJ* 1995;**310**:452–4.
12. Frame J, Evans C, Flaum G, Langford R, Rout P. A comparison of ibuprofen and dihydrocodeine in relieving pain following wisdom teeth removal. *Br Dent J* 1989;**166**:121–4.
13. Galasko C, Russel S, Lloyd J. Double-blind investigation of the efficacy of multiple oral doses of ketorolac tromethamine compared with dihydrocodeine and placebo. *Curr Ther Res* 1989;**45**:844–52.
14. McQuay HJ, Bullingham RE, Moore RA, Carroll D, Evans PJ, O'Sullivan G, *et al.* Zomepirac, dihydrocodeine and placebo compared in postoperative pain after day-case surgery. The relationship between the effects of single and multiple doses. *Br J Anaesth* 1985;**57**:412–19.
15. McQuay HJ, Carroll D, Guest PG, Robson S, Wiffen PJ, Juniper RP. A multiple dose comparison of ibuprofen and dihydrocodeine after third molar surgery. *Br J Oral Maxillofac Surg* 1993;**31**:95–100.

Chapter 14

Dextropropoxyphene in postoperative pain

Summary

The aim of this review was to determine the analgesic efficacy and adverse effects of oral dextropropoxyphene alone and in combination with paracetamol for moderate to severe postoperative pain. Published reports were identified from a variety of electronic databases and additional studies were identified from the reference lists of retrieved reports.

Summed pain intensity and pain relief data were extracted and converted into dichotomous information to yield the number of patients with at least 50% pain relief. This was used to calculate the relative benefit and NNT for one patient to achieve at least 50% pain relief. In six studies (440 patients) dextropropoxyphene was compared with placebo and in five (963 patients) dextropropoxyphene plus paracetamol, 650 mg, was compared with placebo.

For a single dose of dextropropoxyphene, 65 mg, in postoperative pain, the NNT for at least 50% pain relief was 7.7 (95% CI, 4.6–∞) when compared with placebo over 4–6 hours. For the equivalent dose of dextropropoxyphene in combination with paracetamol, 650 mg, the NNT was 4.4 (95% CI, 3.5–5.6) when compared with placebo. Pooled data showed increased incidence of central nervous system adverse effects for dextropropoxyphene plus paracetamol compared with placebo.

Dextropropoxyphene, 65 mg, plus paracetamol, 650 mg, has a similar analgesic efficacy to tramadol, 100 mg, but with a lower incidence of adverse effects. Ibuprofen, 400 mg, has a lower (better) NNT than both dextropropoxyphene, 65 mg, plus paracetamol, 650 mg, and tramadol, 100 mg.

The review has been published in full by Collins and colleagues.[1]

Introduction

Dextropropoxyphene is an opioid analgesic which has been widely available since the 1950s. It is commonly used both alone, and in combination with paracetamol under such brand names as Co-proxamol® and Distalgesic®. In 1996, there were 10 million prescriptions in England for co-proxamol alone, which represents one-fifth of all analgesics prescribed (opiate, non-opiate and NSAIDs); however, it is not clear how much was used for postoperative pain.[2]

Patient surveys have shown that postoperative pain is often not managed well[3] and there is a growing need to assess the efficacy and safety of commonly used analgesics as newer treatments become available. Judging relative analgesic efficacy is difficult as clinical trials use a variety of comparators. It can, however, be determined indirectly by comparing analgesics with placebo in similar clinical circumstances to produce a common analgesic descriptor such as the NNT for at least 50% pain relief. This quantitative systematic review of the analgesic efficacy of dextropropoxyphene has been produced using this method, both with and without paracetamol, allowing comparison with other analgesics.

Methods

A search was undertaken of MEDLINE (1966–November 1996), EMBASE (1980–96), the Cochrane Library (November 1996), Biological Abstracts (1985–96) and the Oxford Pain Relief Database (1950–94)[4] for RCTs of dextropropoxyphene, and its combinations in postoperative pain. The terms 'dextropropoxyphene', 'd-propoxyphene', 'propoxyphene', 'random*', 'clinical trial', 'trial', 'study', analgesi*', 'pain' and 41 brand names (including Distalgesic and Co-proxamol)[5] were used in a broad free text search without restriction to language. Additional reports were identified from reference lists of retrieved articles and reviews. Unpublished data were not sought.

Included reports

The inclusion criteria used were:

- full journal publication
- postoperative pain
- postoperative oral administration
- adult patients
- baseline pain of moderate to severe intensity
- double-blind design

- random allocation to treatment groups which included dextropropoxyphene and placebo or a combination of dextropropoxyphene plus paracetamol and placebo.

Pain outcomes used were TOTPAR or SPID over 4–6 hours or sufficient data provided to allow their calculation. Pain measures allowed for the calculation of TOTPAR or SPID were a standard five-point pain relief scale (none, slight, moderate, good, complete) or a standard four-point pain intensity scale (none, mild, moderate, severe).

Data extraction and analysis

The following were extracted from each study:

- the number of patients treated
- the mean TOTPAR or mean SPID
- study duration
- dose of dextropropoxyphene and paracetamol, where appropriate
- information on adverse effects.

Mean TOTPAR and mean SPID values were converted to % maxTOTPAR or % maxSPID by division into the calculated maximum value.[6] The referenced equations were used to estimate the proportion of patients achieving at least 50% maxTOTPAR.[7,8] The proportions were converted to the number of patients achieving at least 50% maxTOTPAR by multiplying by the total number of patients in the treatment group. The number of patients with at least 50% maxTOTPAR was then used to calculate relative benefit and NNT.

Relative benefit and relative risk estimates with 95% CIs were calculated using the random effects model which provides a more conservative estimate of relative benefit than the fixed effects model.[9] Homogeneity was assumed when $p > 0.1$. A statistically significant benefit of active treatment over placebo was assumed when the lower limit of the 95% CI of the relative benefit was > 1. A statistically significant benefit of placebo over active treatment was assumed when the upper limit of the 95% CI of the relative benefit was < 1. The NNT and NNH with their 95% CIs were calculated.[10] The 95% CI includes no benefit of one treatment over the other when the upper limit is represented as infinity.

Dextropropoxyphene is available as either the hydrochloride or napsylate salt. Equivalent molar doses are 65 mg of dextropropoxyphene hydrochloride and 100 mg of dextropropoxyphene napsylate.

Results

A total of 130 published articles were identified. Two could not be obtained and attempts to contact the authors were unsuccessful. Five citations obtained from reference lists of retrieved reports could not be traced by the British Library. Of the 123 retrieved reports, 33 were not RCTs, 24 were not postoperative pain models or included other pain conditions, 21 were not placebo-controlled, and in five dextropropoxyphene was used as a rescue analgesic only.

Of the 40 RCTs that were placebo-controlled, patients did not have baseline pain of at least moderate severity in ten studies, in 16 there were no pain outcomes which were compatible with our inclusion criteria, and two studies were not double-blind. The data from one study was duplicated and therefore one of the duplicates[11] was excluded. This left 11 reports which met the inclusion criteria and were included in the analysis. Details of the individual studies are presented in *Table 44*.

Dextropropoxyphene versus placebo

Six studies compared dextropropoxyphene hydrochloride, 65 mg (214 patients), with placebo (226 patients), and one trial also compared a dose of 130 mg (25 patients) with placebo (25 patients).

In two trials[12,13] postpartum pain (episiotomy) was investigated, and there were single studies of pain following peridontal surgery,[14] post-urogenital surgery,[15] post-gynaecological surgery,[16] and after various surgical interventions.[17]

The placebo response rate (the proportion of patients experiencing at least 50% pain relief with placebo) varied between 4% and 76%. The dextropropoxyphene response rate (the proportion of patients experiencing at least 50% pain relief with dextropropoxyphene) varied between 19% and 84% *(Figure 39)*. Data were homogenous ($p = 0.13$). Dextropropoxyphene, 65 mg, was not significantly different from placebo, relative benefit 1.4 (95% CI, 0.97–2.0) *(Table 45)*.

For a single dose of dextropropoxyphene, 65 mg, the NNT was 7.7 (95% CI, 4.6–∞) for at least 50% pain relief over a period of 4–6 hours compared with placebo for pain of moderate to severe intensity.

Pooled relative benefit estimates were calculated using the random effects model.[9]

One trial[17] used a dose of 130 mg of dextropropoxyphene (25 patients). The relative benefit

TABLE 44 *Dextropropoxyphene in postoperative pain: details of included studies*

Study	Condition and number of patients	Study design, duration and follow-up	Outcome measures	Dosing regimen	Analgesic outcome results	Remedication	Withdrawals and exclusions	Adverse effects
Dextropropoxyphene plus paracetamol								
Cooper, *et al.*, 1981[19]	Dental surgery n = 248 Age: adult	RCT, double-blind, single oral dose, parallel groups, general or local anaesthetic Self-assessed at home at 0, 1 hour then hourly for 4 hours. Medication given when pain of moderate to severe intensity	PI (4-point scale) standard PR (5-point scale) standard Global evaluation by patient (5-point scale) at 4 hours	Dextropropoxyphene napsylate, 100 mg, + paracetamol, 650 mg (n = 42); placebo (n = 37)	Combination of dextropropoxyphene with paracetamol significantly better than placebo for SPID and TOTPAR ($p < 0.001$). 4-hour TOTPAR: dextropropoxyphene + paracetamol 8.31, placebo 3.38.	Allowed at > 1 hour; if remedicated before patient withdrawn from study. If remedicated, PR recorded as 0, and last PI score prior to remedication taken for all further time points.	200 analysed, 48 excluded: 31 violated protocol, 17 did not take medication.	None serious and no patients withdrew as result. Dextropropoxyphene + paracetamol, 5/42 with 5 adverse effects; placebo, 4/37 with 5.
Cooper, 1980[18]	Dental surgery n = 179 Age: adult	RCT, double-blind, single oral dose, parallel groups, mostly local anaesthetic. Self-assessed at 0, 1 hour then hourly for 4 hours. Medication given when pain of moderate to severe intensity.	PI (4-point scale) standard PR (5-point scale) standard Global evaluation by patient (5-point scale) at 4 hours	Dextropropoxyphene napsylate, 100 mg, + paracetamol, 650 mg (n = 40); placebo (n = 48)	Combination of dextropropoxyphene with paracetamol significantly better than placebo for SPID and TOTPAR ($p < 0.05$). 4-hour TOTPAR: dextropropoxyphene + paracetamol 5.65, placebo 4.17.	Did not state when remedication allowed. If remedicated, last PR and PI score before remedication used for all further time points.	179 analysed. No withdrawals reported.	None serious and no patients withdrew as a result. Dextropropoxyphene + paracetamol, 10/40 with 13 adverse events. Placebo: 13/48 with 17 adverse events.
Evans, *et al.*, 1982[20]	Minor orthopaedic surgery n = 120 Age: adult	RCT, double-blind, single oral dose, parallel groups, general anaesthetic. Assessed by same nurse observer at 0, 0.5, 1 hour then hourly for 4 hours. Medication given when pain of moderate to severe intensity.	PI (4-point scale) standard PR (5-point scale) standard	Dextropropoxyphene hydrochloride, 65 mg, + paracetamol, 650 mg (n = 30); placebo (n = 30)	Dextropropoxyphene + paracetamol significantly better than placebo ($p < 0.05$) for TOTPAR. 4-hour TOTPAR: dextropropoxyphene + paracetamol 7.37, placebo 4.70.	If remedicated before 4 hours, last PI and PR score prior to remedication used for all further time points.	120 analysed. No withdrawals were reported.	None serious and no patients withdrew as a result. Dextropropoxyphene + paracetamol: 16/30 with 16 adverse events. Placebo: 13/30 with 13 adverse events.
Honig & Murray, 1981[21]	Postoperative, primarily orthopaedic n = 196 Age range: 19–74 years	RCT, double-blind, single oral dose, parallel groups. Assessed by nurse observer at 0, 0.5, 1 hour then hourly for 6 hours. Medication given when pain of moderate to severe intensity.	PI (4-point scale) standard PR (5-point scale) non-standard Global evaluation by patient at 5 hours (5-point)	Dextropropoxyphene napsylate, 100 mg, + paracetamol, 650 mg (n = 50); placebo (n = 48)	Combination of dextropropoxyphene with paracetamol significantly better than placebo ($p < 0.05$) for SPID and TOTPAR. 6-hour TOTPAR: dextropropoxyphene + paracetamol 8.04, placebo 5.49.	If remedicated within 6 hours patient's overall rating of drug taken at time of remedication.	196 analysed. No withdrawals reported.	Authors did not give details of adverse events but reported NSD between active drug and placebo groups.
Moore & McQuay, 1997[7]	Dental + postoperative pain n = 638 Age: adult	Individual patient data from 18 double-blind, RCTs. Study duration 8 hours. Single oral dose, parallel groups. Medication was given when pain of moderate to severe intensity.	Number of patients with at least 50% maxTOTPAR	Dextropropoxyphene napsylate, 100 mg, + paracetamol, 650 mg (n = 316); placebo (n = 322)	Number of patients with at least 50% maxTOTPAR: dextropropoxyphene napsylate, 100 mg, + paracetamol, 650 mg, 112/316; placebo, 41/322.	None reported.	None reported.	None serious and no withdrawals. Dextropropoxyphene + paracetamol, 88/316 adverse events; placebo, 66/322 adverse events. Significantly higher incidence with active treatment than placebo for: dizziness, relative risk 2.0 (1.1–4.0); drowsiness/ somnolence, 2.16 (1.5–3.2).
Dextropropoxyphene alone								
Berry, *et al.*, 1975[12]	Postpartum pain (episiotomy) n = 225 Age range: 15–39 years	RCT, double-blind, single oral dose, parallel groups. Assessed by observer in hospital at 0.5, 1 hour, then hourly for 4 hours. Medication taken when pain of moderate to severe intensity.	PI (4-point scale) standard PR (5-point) non-standard Global rating (5-point scale) by patient	Dextropropoxyphene hydrochloride, 65 mg (n = 730; placebo (n = 76)	Dextropropoxyphene significantly better than placebo ($p < 0.01$). Global evaluation (good or excellent PR): dextropropoxyphene 26/73, placebo 18/76.	Patients allowed to remedicate after reasonable amount of time. If remedicated, patients were regarded as a treatment failure.	225 analysed; no details given.	No adverse effects were reported with either active treatment or placebo.

continued

TABLE 44 contd *Dextropropoxyphene in postoperative pain: details of included studies*

Study	Condition and number of patients	Study design, duration and follow-up	Outcome measures	Dosing regimen	Analgesic outcome results	Remedication	Withdrawals and exclusions	Adverse effects
Dextropropoxyphene alone contd								
Bloomfield *et al.*, 1980[13]	Postpartum pain (episiotomy) n = 100 Age: adult	RCT, double-blind, single oral dose, parallel groups. Assessed, in hospital, by same nurse observer at 0, 0.5, 1 hour then hourly for 6 hours. Medication taken when pain of moderate to severe intensity.	PI (4-point scale) standard PR not measured	Dextropropoxyphene hydrochloride, 65 mg (n = 25); placebo (n = 25)	Dextropropoxyphene not significantly better than placebo at 10% probability level. SPID at 6 hours: dextropropoxyphene 9.32, placebo 8.12.	If remedicated patients withdrawn from study. Subsequent PR readings set to pre-treatment score.	100 analysed. 6 withdrew: no pain relief or patients remedicated.	No serious adverse effects were reported and no patients withdrew as a result. Dextropropoxyphene, 6/25 with 12 adverse events; placebo, 9/25 with 9 adverse events.
Cooper, *et al.*, 1986[14]	Peridontal surgery n = 301 Age: adult	RCT, double-blind, single oral dose, parallel groups, local anaesthetic. Self-assessed at 0, 0.5, 1 hour then hourly for 6 hours. Medication taken when pain of moderate to severe intensity.	PI (4-point scale) standard PR (5-point scale) standard Global evaluation by patient at 6 hours (5-point)	Dextropropoxyphene hydro chloride, 65 mg (n = 50); placebo (n = 56)	Dextropropoxyphene significantly better than placebo ($p < 0.1$). TOTPAR at 6 hours: dextropropoxyphene 7.7, placebo 5.2.	Allowed after 1 hour. Last score prior to remedication was used for duration of study.	212 analysed, 91 excluded: 48 did not medicate, 17 missed readings, 9 lost to follow-up, 4 remedicated at < 1 hour, 3 remedicated with slight pain, 4 uninterpretable data, 2 took other medication, 2 did not receive study medicine, 1 lost form.	No serious adverse effects were reported and no patients withdrew as a result. Dextropropoxyphene, 10/50 with 10 adverse events; placebo, 5/56 with 5 adverse events.
Coutinho, et al., 1976[15]	Urogenital surgery n = 15 Age: adult	RCT, double-blind, single oral dose, parallel groups, local anaesthetic. Assessed by observer at 0, 0.5, 1 hour then hourly for 5 hours. Medication taken when pain of moderate to severe intensity.	PI (4-point scale) standard PR (5-point scale) non-standard	Dextropropoxyphene hydrochloride, 65 mg (n = 15); placebo (n = 15)	Dextropropoxyphene not significantly better than placebo (*p* not given). Mean SPID at 5 hours: dextropropoxyphene 4.5, placebo 3.3.	Allowed at 4 hours if no PR. If remedicated before 4 hours, patients withdrawn from study.	No exclusions or withdrawals.	No serious adverse effects were reported and no patients withdrew as a result. Dextropropoxyphene, 1/15 with 1 adverse event; placebo, 0/15.
Trop, *et al.*, 1979[17]	Postoperative pain, various procedures n = 125 Age range: 18–73 years	RCT, double-blind, single oral dose, parallel groups, local anaesthetic. Assessed by observer at 0, 0.5, 1 hour then hourly for 5 hours. Medication taken when pain of moderate to severe intensity.	PI (4-point scale) standard PR (5-point scale) standard	Dextropropoxyphene hydro chloride, 65 mg (n = 25); dextropropoxyphene hydrochloride, 130 mg (n = 25); placebo (n = 25)	Dextropropoxyphene, 130 mg, significantly better than placebo ($p < 0.01$). SPID and TOTPAR given at 6 hours. TOTPAR: dextropropoxyphene, 65 mg, 8.54; dextropropoxyphene, 130 mg, 9.03; placebo, 2.68.	Did not state minimum time allowed for remedication. Last PR score before remedication used for all further time points.	78 patients analysed. 47 excluded due to protocol violation.	Authors reported significant difference from placebo for central nervous system adverse events ($p = 0.05$). None serious and no withdrawals. Dextropropoxyphene hydrochloride, 65 mg, 19/25 with 27 adverse events; dextropropoxyphene hydrochloride, 130 mg, 23/25 with 34 adverse events; placebo, 10/25 with 12 adverse events.
van Staden, 1971[16]	Gynaecological surgery n = 91 Age: adult	RCT, double-blind, crossover design, general anaesthetic. Self-assessed at 1 hour then hourly for 8 hours. Medication given when pain of moderate to severe intensity.	PI (4-point scale) standard PR measured as PI difference	Dextropropoxyphene hydrochloride, 65 mg (n = 26); placebo (n = 29)	Dextropropoxyphene not significantly better than placebo (*p* not given). SPID at 4 hours: dextropropoxyphene hydrochloride, 65 mg, 1.64; placebo, 1.57.	Allowed after 1 hour if no PR. PR scored as zero for all subsequent time points.	80 patients analysed, 11 excluded: 6 violated protocol, 2 vomited, 3 had insufficient pain.	No serious adverse effects were reported and no patients withdrew as a result. Dextropropoxyphene, 4/26 with 4 adverse events; placebo, 1/29.

estimate for dextropropoxyphene, 130 mg, compared with placebo was 10 (95% CI, 1.4–73). The NNT was 2.8 (95% CI, 1.8–6.5) for at least 50% pain relief over a period of 5 hours compared with placebo for pain of moderate to severe intensity.

Adverse effects

Details of adverse effects are presented in *Table 46*. No patients withdrew as a result of adverse effects and all were reported as transient and of mild to moderate severity. One study reported no adverse effects with either placebo or active treatment.[12]

In one study, the authors reported that dextropropoxyphene, both 65 mg and 130 mg, have a significantly higher incidence of grogginess, sleepiness, and lightheadedness than placebo ($p = 0.05$).[17]

However, pooled data from the four trials reporting either drowsiness, sleepiness or somnolence[13–15,17] showed no significant difference in incidence between dextropropoxyphene, 65 mg, (18/115) and placebo (15/121), with a relative risk of 1.3 (95% CI, 0.7–24). No other trial reported lightheadedness or grogginess in the dextropropoxyphene group.

Dextropropoxyphene plus paracetamol versus placebo

In four reports dextropropoxyphene napsylate, 100 mg, plus paracetamol, 650 mg, was compared with placebo, and in one dextropropoxyphene hydrochloride, 65 mg, plus paracetamol, 650 mg. A total of 478 patients received dextropropoxyphene plus paracetamol, and 485 patients received placebo.

One report[7] was a meta-analysis of individual patient data from 18 studies with dichotomous information (the number of patients achieving at least 50% maxTOTPAR); eight reports investigated dextropropoxyphene napsylate, 100 mg, plus paracetamol, 650 mg. Only one of the studies had been published and the duplicate publication was excluded.[11]

In two reports[18,19] pain following dental surgery (impacted third molar) was studied, in two others[20,21] pain post orthopaedic surgery, and in one report[7] pain following both dental and general surgery (abdominal, orthopaedic and gynaecological) was studied.

The placebo response rate varied between 6% and 27%. The dextropropoxyphene plus paracetamol response rate varied between 25% and 57% *(Figure 39)*. The trial results were homogenous ($p = 0.35$). Dextropropoxyphene (hydrochloride, 65 mg, or napsylate, 100 mg) plus paracetamol, 650 mg, was significantly superior to placebo, relative benefit 2.4 (95% CI, 1.9–3.1) *(Table 45)*. For a single dose of dextropropoxyphene (hydrochloride, 65 mg, or napsylate, 100 mg) plus paracetamol, 650 mg, the NNT was 4.4 (95% CI, 3.5–5.6) for at least 50% pain relief over 4–6 hours compared with placebo for pain of moderate to severe intensity.

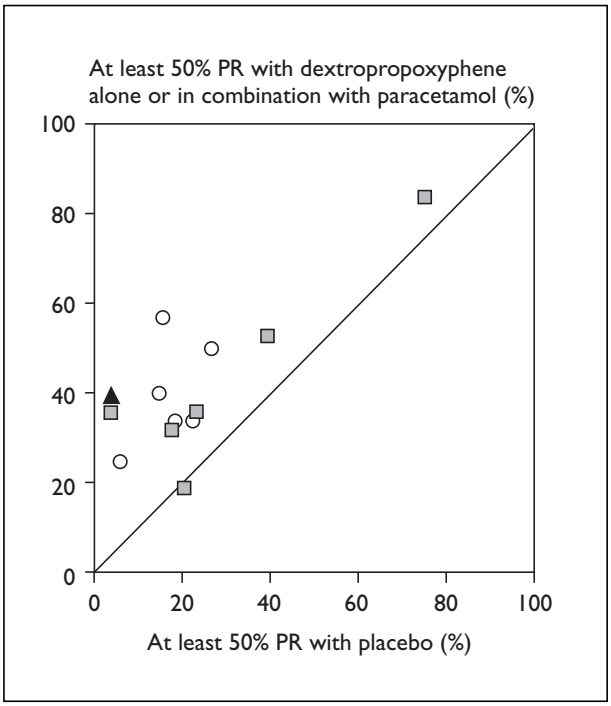


FIGURE 39 *Trials of oral dextropropoxyphene in postoperative pain (■, dextropropoxyphene HCl, 65 mg; ○, dextropropoxyphene (napsylate, 100 mg, or HCl, 65 mg) + paracetamol, 650 mg; ▲, dextropropoxyphene HCl, 130 mg;)*

Adverse effects

Details of adverse effects are given in *Table 46*. No patients withdrew as a result of adverse effects and all were reported as transient and of mild to moderate severity. In one trial,[21] details of adverse effects were not given but it was reported that there was no significant difference between active and placebo groups. The individual patient meta-

TABLE 45 *Summary of relative benefit and NNT for trials of dextropropoxyphene and dextropropoxyphene plus paracetamol against placebo*

Number of trials	Dose of dextropropoxyphene	Number of patients with > 50% PR: dextropropoxyphene	Number of patients with > 50% PR: placebo	RB (95% CI)	NNT (95% CI)
Dextropropoxyphene alone against placebo					
6	65 mg	85/214	60/226	1.4 (0.97–2.0)	7.7 (4.6–∞)
1	130 mg	10/25	1/25	10.0 (1.4–73)	2.8 (1.8–6.5)
Dextropropoxyphene plus paracetamol against placebo					
6	65 mg hydrochloride or 100 mg napsylate	184/478	74/485	2.4 (1.9–3.1)	4.4 (3.5–5.6)

TABLE 46 *Summary of adverse effects for trials of dextropropoxyphene and dextropropoxyphene plus paracetamol against placebo*

Number of trials	Adverse events	Number of patients with adverse events: drug	Number of patients with adverse events: placebo	RR (95% CI)	NNH (95% CI)
Dextropropoxyphene					
2	Nausea	3/75	2/81	1.6 (0.3–9.4)	N/A
3	Drowsiness/sleepiness/ somnolence	18/115	15/121	1.3 (0.7–2.4)	N/A
2	Headache	5/75	3/81	1.6 (0.5–4.9)	N/A
Dextropropoxyphene plus paracetamol					
3	Nausea	12/405	33/799	0.7 (0.4–1.4)	N/A
1	Vomiting	2/323	6/714	1.4 (0.3–6.7)	N/A
4	Dizziness	17/435	16/829	2.2 (1.1–4.3)	50 (24–∞)
3	Drowsiness/somnolence	57/405	55/799	2.2 (2.0–2.4)	14 (9.1–30)
4	Headache	14/435	51/829	0.5 (0.4–0.6)	–33 (–170, –19)

Negative NNTs indicate that fewer headaches occur with dextroropoxyphene plus paracetamol than with placebo.

analysis[7] pooled data on adverse effects from all 18 placebo groups; 714 patients received placebo.

The incidence of drowsiness or somnolence was reported in three studies.[7,18,19] The pooled data indicated a significantly higher incidence in the dextropropoxyphene combination group (57/405) than in the placebo group (55/799), with a relative risk of 2.2 (95% CI, 2.0–2.4) and an NNH of 14 (95% CI, 9.1–30).

Dizziness was reported in four trials.[7,18–20] Pooled data indicated a significantly higher incidence of dizziness with dextropropoxyphene plus paracetamol (17/435) than with placebo (16/829), with a relative risk of 2.2 (95% CI, 1.1–4.3) and an NNH of 50 (95% CI, 24–∞).

The incidence of headache was reported in four trials.[18–21] The pooled data showed dextropropoxyphene plus paracetamol (14/435) to have a significantly lower incidence of headache than placebo (51/829), with a relative risk of 0.5 (95% CI, 0.4–0.6) and an NNH of –33 (95% CI, –170, –19).

Three trials considered the incidence of nausea.[7,18,19] Pooled data showed no significant difference with dextropropoxyphene plus paracetamol (12/405) than with placebo (33/799), relative risk 0.7 (95% CI, 0.4–1.4).

Incidence of vomiting was reported in one study.[7] For dextropropoxyphene plus paracetamol (2/323) it was not significantly different from placebo (6/714), relative risk 1.4 (95% CI, 0.3–6.7).

Comment

For a single dose of dextropropoxyphene, 65 mg, the NNT was 7.7 (95% CI, 4.6–∞) for at least 50% pain relief compared with placebo. This means that one in every eight patients with pain of moderate to severe intensity would experience at least 50% pain relief with dextropropoxyphene hydrochloride, 65 mg, who would not have done so with placebo. The 95% CI included no benefit. The equivalent NNT for a single dose of dextropropoxyphene (65 mg, hydrochloride or 100 mg, napsylate) plus paracetamol, 650 mg, was 4.4 (95% CI, 3.5–5.6), indicating higher efficacy. The 95% CIs of dextropropoxyphene alone and the combination with paracetamol overlapped. The dextropropoxyphene/paracetamol combination had an NNT similar to that of both paracetamol, 1000 mg, and ibuprofen, 200 mg. Both ibuprofen, 400 mg, and diclofenac, 50 mg, had NNTs with 95% CIs lower (better) than that of the combination and which did not overlap with it.

For a single dose of dextropropoxyphene, 130 mg, the NNT was 2.8 (95% CI, 1.8–6.5). This appears to show a dose response for dextropropoxyphene. However, given the overlapping CIs and the very small number of patients in the dextropropoxyphene, 130 mg, trial,[50] this conclusion is not robust.

A single dose of dextropropoxyphene, 65 mg, plus paracetamol, 650 mg, showed a significantly higher incidence of central nervous system adverse effects (somnolence, dizziness) than placebo *(Table 46)*.

These adverse effects have also been shown for tramadol, 100 mg, with lower (worse) NNHs for both dizziness and somnolence.[7] Tramadol, 100 mg, also showed a significantly higher incidence of nausea and vomiting than placebo. These adverse effects were reported with dextropropoxyphene, 65 mg, plus paracetamol, 650 mg, but the incidence was not significantly different from placebo.

The combination of dextropropoxyphene, 65 mg, with paracetamol, 650 mg, showed similar efficacy to tramadol, 100 mg, for single dose studies in postoperative pain with a lower incidence of adverse effects.

References

1. Collins SL, Edwards J, Moore A, McQuay HJ. Oral dextropropoxyphene in postoperative pain: a quantitative systematic review. *Eur J Clin Pharmacol* 1998;**54**:107–12.
2. Government Statistical Service. Prescription cost analysis for England 1995. London: Department of Health, 1996.
3. Bruster S, Jarman B, Bosanquet N, Weston D, Erens R, Delbanco TL. National survey of hospital patients. *BMJ* 1994;**309**:1542–6.
4. Jadad AR, Carroll D, Moore A, McQuay H. Developing a database of published reports of randomised clinical trials in pain research. *Pain* 1996;66:239–46.
5. Reynolds JEF. Martindale: the extra pharmacopoeia. 30th edition. London: Pharmaceutical Press, 1993:292–314.
6. Cooper SA. Single-dose analgesic studies: the upside and downside of assay sensitivity. *Adv Pain Res Ther* 1991;**18**:117–24.
7. Moore RA, McQuay HJ. Single-patient data meta-analysis of 3453 postoperative patients: oral tramadol versus placebo, codeine and combination analgesics. *Pain* 1997;**69**:287–94.
8. Moore A, Moore O, McQuay H, Gavaghan D. Deriving dichotomous outcome measures from continuous data in randomised controlled trials of analgesics: use of pain intensity and visual analogue scales. *Pain* 1997;**69**:311–15.
9. DerSimonian R, Laird N. Meta-analysis of clinical trials. *Control Clin Trials* 1986;**7**:177–88.
10. Cook RJ, Sackett DL. The number needed to treat: a clinically useful measure of treatment effect. *BMJ* 1995;**310**:452–4.
11. Sunshine A, Olson NZ, Zighelboim I, DeCastro A, Minn FL. Analgesic oral efficacy of tramadol hydrochloride in postoperative pain. *Clin Pharmacol Ther* 1992;**51**:740–6.
12. Berry FN, Miller JM, Levin HM, Bare WW, Hopkinson JH III, Feldman AJ. Relief of severe pain with acetaminophen in a new dose formulation versus propoxyphene hydrochloride 65 mg and placebo: a comparative double blind study. *Curr Ther Res* 1975;**17**:361–8.
13. Bloomfield SS, Barden TP, Mitchell J. Nefopam and propoxyphene in episiotomy pain. *Pharmacol Ther* 1980;**27**:502–7.
14. Cooper SA, Wagenberg B, Zissu J, Kruger GO, Reynolds DC, Gallegos LT, *et al.* The analgesic efficacy of suprofen in peridontal and oral surgical pain. *Pharmacotherapy* 1986;**6**:267–76.
15. Coutinho A, Bonelli J, Nanci de Carvelho P. A double-blind study of the analgesic effects of fenbufen, codeine, aspirin, propoxyphene and placebo. *Curr Ther Res* 1976;**19**:58–65.
16. van Staden MJ. The use of gilfanan in postoperative pain. *S Afr Med J* 1971;**45**:1235–7.
17. Trop D, Kenny L, Grad BR. Comparison of nefopam hydrochloride and propoxyphene hydrochloride in the treatment of postoperative pain. *Can Anaesth Soc J* 1979;**26**:296–304.
18. Cooper SA. Double-blind comparison of zomepirac sodium, propoxyphene/acetaminophen, and placebo in the treatment of oral surgical pain. *Curr Ther Res* 1980;**28**:630–8.
19. Cooper SA, Breen JF, Giuliani RL. The relative efficacy of indoprofen compared with opioid-analgesic combinations. *J Oral Surg* 1981;**39**:21–5.
20. Evans PJ, McQuay HJ, Rolfe M, O'Sullivan G, Bullingham RE, Moore RA. Zomepirac, placebo and paracetamol/dextropropoxyphene combination compared in orthopaedic postoperative pain. *Br J Anaesth* 1982;**54**:927–33.
21. Honig S, Murray KA. Surgical pain: zomepirac sodium, propoxyphene/acetaminophen combination, and placebo. *J Clin Pharmacol* 1981;**21**:443–8.

Chapter 15

Oral tramadol, codeine and combination analgesics in postoperative pain

Summary

The analgesic effectiveness and safety of oral tramadol was compared with standard analgesics using a meta-analysis of individual patient data from RCTs in patients with moderate or severe pain after surgery or dental extraction. Calculation of % maxTOTPAR from individual patient data, and the use of at least 50% maxTOTPAR defined clinically acceptable pain relief. NNT for one patient to have at least 50% maxTOTPAR compared with placebo was used to examine the effectiveness of different single oral doses of tramadol and comparator drugs.

A total of 18 randomised, double-blind, parallel group single-dose trials with 3453 patients using categorical pain relief scales allowed the calculation of % maxTOTPAR. The use of at least 50% maxTOTPAR was a sensitive measure to discriminate between analgesics.

Tramadol and comparator drugs gave significantly more analgesia than placebo. In postsurgical pain, tramadol, 50, 100 and 150 mg, had NNTs for at least 50% maxTOTPAR of 7.3 (95% CI, 4.6–18), 4.8 (3.4–8.2) and 2.4 (2.0–3.1), respectively, comparable with aspirin, 650 mg, plus codeine, 60 mg (NNT 3.6 (95% CI, 2.5–6.3)) and paracetamol, 650 mg, plus propoxyphene, 100 mg (NNT 4.0 (3.0–5.7)). With the same dose of drug, postsurgical patients had more pain relief than those having dental surgery. Tramadol showed a dose response for analgesia in patients with both postsurgical and dental pain.

With the same dose of drug, postsurgical pain patients had fewer adverse events than those having dental surgery. Adverse events – headache, nausea, vomiting, dizziness, somnolence – for tramadol, 50 mg and 100 mg, had a similar incidence to comparator drugs. There was a dose response with tramadol, tending towards higher incidences at higher doses.

Single patient meta-analysis using more than half pain relief provides a sensitive description of the analgesic properties of a drug, and NNT calculations allow comparisons to be made with standard analgesics.

This chapter of the review has been published in full by Moore and McQuay.[1]

Introduction

Study of analgesics still poses problems, some 40 years after Beecher first described methods of measuring pain and pain relief.[2,3] These can be of different sorts, starting with the obvious but important: the many possible comparisons of drugs, doses, routes of administration and pain condition which makes meaningful comparison difficult. Many controlled trials have been performed and many published; some 10,000 RCTs (over 4000 in pharmacological interventions in acute pain) have been identified.[4]

Quantitative systematic reviews pool data from a number of trials; while individual trials may have relatively small numbers of patients receiving a particular treatment, meta-analysis allows the result to be confirmed using data from many patients in many trials, thereby increasing the power to determine the 'true' result. It can therefore provide a higher quality of evidence on which to base decisions by prescribers, policy-makers and patients.

Choice of RCTs for systematic reviews is essential. Randomisation (and concealment) of treatment allocation limits selection bias, and blinding of treatments controls observer bias. Inadequacies of randomisation or blinding exaggerate estimates of treatment effect.[5] In RCTs in pain relief, standard methods of measuring pain relief and trial conduct appear to be effectively blinded.[6]

Results of systematic reviews have to be easily understood to be useful and used. The elegant NNT approach[7] involves defining a clinical endpoint, and comparing the rate of that event in a treatment group with the rate in a comparator group; NNT calculations require dichotomous data. NNTs derived for particular benefits or harm can provide a useful starting point for simple verbal

and numerical results accessible to any doctor or patient.[8]

Meta-analysis using individual patient data sometimes produces lower estimates of effect of treatment than does meta-analysis using group descriptions,[9] although the generality of this has been challenged.[10] Using individual patient information may not always be possible but where possible it is preferred[11] because it is claimed to have the least bias of any meta-analytical method.

Studies of pain relief may present an additional complication. The classic design of single dose oral medication with both placebo and active controls to demonstrate analgesic sensitivity[3] is explanatory.[12] Such trials provide evidence that a drug is an analgesic rather than information about the best way to use the compound in practice. In this context, NNT methods are useful indicators of relative efficacy.[13]

Tramadol has been used in many European countries since the late 1970s, in many different pain conditions. Since most studies with tramadol in Europe had not been conducted according to US regulatory requirements, a completely new programme of clinical studies for registration in the USA of an oral tramadol formulation took place in the late 1980s and early 1990s. A total of 18 single-dose studies were conducted, nine in dental pain models and nine in postsurgical pain, and the results of the studies have been summarised.[14]

We performed a single-patient data meta-analysis of these 18 studies, and any others (published or unpublished) which could be found and which had categorical pain relief scales, allowing the calculation of the percentage of maximum pain relief obtained by individual patients. Combining data from many studies will help shed additional light on the debate[15,16] about the conflicting results of tramadol in postoperative pain.[17,18]

Methods

Primary trials

Individual patient data from 18 primary trials were made available by Grünenthal GmbH, Aachen, Germany and Robert Wood Johnson Pharmaceutical Research Institute, Spring House, Pennsylvania, USA. One of these studies had been published.[17] Other studies which used single doses of oral tramadol with categorical pain relief scoring in acute painful conditions were sought by reference to the in-house data from Grünenthal GmbH, from Searle (UK) Ltd, and by searching MEDLINE (1960–95) and the Oxford Pain Relief Database (1950–95)[4] using tramadol as a free-text term.

There was a prior hypothesis that analgesic drugs may produce different analgesic responses in painful dental procedures (such as third molar extractions) than postsurgical procedures (such as abdominal, orthopaedic or gynaecological operations). The prior intention, therefore, was to analyse these conditions separately. Included reports were scored for inclusion and methodological quality using a 3-item scale.[19]

Protocols for the RWJ studies of postsurgical pain and of pain due to the extraction of impacted third molars were essentially identical. Trials were of double-blind, single-dose, parallel-group design; randomisation was by computerised random-number generation, stratified on pre-treatment pain intensity. Criteria for patient selection were moderate or severe pain and that the patient's condition was appropriate for management with a centrally acting analgesic and paracetamol. The age range was from 18 to 70 years. Patients had to be cooperative, reliable and motivated, and be able to take oral medication. Exclusion criteria included patients with mild or no pain, those who had taken analgesic drugs within 3 hours of study drug administration, those needing sedatives during the observation period and those with known contraindications or medical conditions which might interfere with observations.

Drugs were given as single oral doses: placebo (695 evaluable patients); codeine, 60 mg (649); tramadol, 50 mg (409); tramadol, 75 mg (281); tramadol, 100 mg (468); tramadol, 150 mg (279); tramadol, 200 mg (50); aspirin, 650 mg, plus codeine, 60 mg (305); paracetamol, 650 mg, plus propoxyphene, 100 mg (316).

Patients were given the study drug if they had moderate or severe pain on a four-point categorical scale (0 = no pain, 1 = slight, 2 = moderate, 3 = severe). Thereafter observations were made at 30 minutes, and 1, 2, 3, 4, 5 and 6 hours after administration. Pain intensity was measured using the same categorical scale, together with a five-point CATPR (0 = no relief, 1 = a little, 2 = some, 3 = a lot, 4 = complete). Time of remedication was also recorded, as well as a global assessment of therapy (excellent, very good, good, fair or poor) at the final evaluation.

Adverse experiences volunteered by the patient after non-directive questioning were recorded regardless of any rescue medication used.

Calculations

For each patient the area under the curve of pain relief (categorical scale) against time was calculated (TOTPAR) for 6 hours after the study drug was given. If patients remedicated, pain relief scores reverted to zero and pain intensity scores to the initial value; adverse event recording but not pain evaluations continued after remedication. The percentage of the maximum possible for this summary measure was then calculated (% maxTOTPAR).[20] The number of patients on each treatment who achieved more than 50% maxTOTPAR was determined.

Relative benefit (which indicates how much more likely is an individual given a particular treatment to have a specific outcome than someone not given the treatment) and its 95% CIs were calculated for individual trials using a fixed-effects model,[21] and NNT using the method of Cook and Sackett.[7] The same method was used to calculate the NNH for adverse effects. Relative risk and NNT are given with 95% CIs in text and tables. Significance testing for dose response of tramadol was performed using the Kruskal–Wallis test unstratified for type of surgery.

Results

Search results

Individual patient data for 3453 patients from 18 studies was supplied. It consisted of pain intensity and pain relief scores from start of study to 8 hours post dose and aggregate adverse effect information. Studies with their codes, drug treatments and numbers of patients, are presented in *Table 47*. Data on pain measurements and adverse effects for these single-dose parallel-group double-blind studies was provided. Of the nine post-surgical pain studies, two (TR and TV, see *Table 47)* followed Caesarean section, and one (TX) was conducted with outpatients. Of the nine dental pain studies, three (TI, TI2 and TO) were conducted with outpatients. Tramadol, 200 mg, was given in only one study in dental pain and these data were excluded. Study reports were of high methodological quality, scoring the maximum of 5 points on a validated scale.[19]

Literature searches through MEDLINE found two relevant studies of oral tramadol in postoperative pain.[17,18] The first[17] formed part of the data set supplied. The other,[18] which did not show a significant difference between tramadol, 50 and 100 mg, and placebo, used a pain intensity scoring system rather than pain relief, and therefore had to be excluded from this analysis. Another study[22] using several dose levels of oral tramadol after dental surgery was multiple dose and used only pain intensity scoring. It showed significant differences between all tramadol doses and placebo, but again could not be used. No other relevant studies were identified.

Analgesic efficacy

The relative benefits and NNTs for each drug tested are shown in *Table 48,* for dental and post-operative pain both separately and combined. The proportions of patients achieving at least 50% maxTOTPAR are shown in *Figure 40*. There was a clear dose response for tramadol ($p < 0.0001$, Kruskal–Wallis test).

Dental pain

Among the dental studies, all treatments showed significantly greater pain relief (greater proportion of patients with at least 50% of % maxTOTPAR) than with placebo (relative risk lower CI > 1) except for codeine, 60 mg. There was a clear dose response for tramadol, with higher odds ratios and lower NNT values with the higher doses. Tramadol, 100 mg and 150 mg, produced NNT values of 4.6 (95% CI, 3.6–6.4) and 4.1 (95% CI, 2.9–7.3) respectively, lower than aspirin/codeine (NNT 6.3 (95% CI, 4.5–9.8)) and paracetamol/propoxyphene (NNT 5.3 (95% CI, 3.4–11.4)).

Postsurgical pain

All treatments showed statistically significantly superior analgesia to placebo. There was a clear dose response for tramadol; tramadol, 100 mg, had an NNT of 4.8 (95% CI, 3.4–8.2) and tramadol, 150 mg, an NNT of 2.4 (95% CI, 2.0–3.1). This was lower than aspirin/codeine and paracetamol/propoxyphene combinations with NNT values of 3.5 (95% CI, 2.5–6.3) and 4.0 (95% CI, 3.0–5.7), respectively.

Dental and postsurgical pain models compared

With the exception of tramadol, 100 mg, NNTs were lower in postsurgical pain than in dental pain. When the numbers of patients with more than 50% pain relief were compared for each treatment between postsurgical and dental pain *(Figure 40),* some treatments produced significantly more pain relief in postsurgical pain. This was the case for codeine, 60 mg (relative benefit 2.4 (95% CI, 1.6–3.5)), tramadol, 75 mg (2.4 (95% CI, 1.5–3.8)) and 150 mg (1.9 (95% CI, 1.4–2.6)), aspirin plus codeine (1.6 (95% CI, 1.04–2.3)) and paracetamol plus propoxyphene (1.6 (95% CI, 1.1–2.4)).

TABLE 47 *Oral tramadol, codeine and combination analgesics in postoperative pain: trials, treatments and patient numbers*

Postsurgical			Dental		
Trial	**Drug**	**Patients**	**Trial**	**Drug**	**Patients**
TA	Tramadol, 50 mg	52	TE	Tramadol, 50 mg	28
	Tramadol, 100 mg	58		Tramadol, 100 mg	30
	Codeine, 60 mg	26		ASA 650 & C60	28
	Placebo	28		Codeine, 60 mg	28
				Placebo	29
TC	Tramadol, 50 mg	40	TE2	Tramadol, 50 mg	23
	Tramadol, 100 mg	39		Tramadol, 100 mg	21
	ASA 650 & C 60	40		ASA 650 & C60	22
	Codeine, 60 mg	39		Codeine, 60 mg	24
	Placebo	40		Placebo	21
TR	Tramadol, 75 mg	40	TF	Tramadol, 50 mg	47
	Tramadol, 150 mg	40		Tramadol, 100 mg	49
	APAP 650 & P 100	41		ASA 650 & C60	45
	Placebo	40		Codeine, 60 mg	47
				Placebo	49
TV	Tramadol, 50 mg	31	TG	Tramadol, 50 mg	50
	Tramadol, 100 mg	31		Tramadol, 100 mg	49
	ASA 650 & C 60	30		ASA 650 & C60	41
	Codeine, 60 mg	29		Codeine, 60 mg	33
	Placebo	28		Placebo	27
TW	Tramadol, 50 mg	40	TH	Tramadol, 50 mg	47
	Tramadol, 100 mg	40		Tramadol, 100 mg	51
	APAP 650 & P100	39		ASA 650 & C60	51
	Codeine, 60 mg	39		Codeine, 60 mg	48
	Placebo	40		Placebo	50
TW2	Tramadol, 75 mg	41	TI	Tramadol, 50 mg	51
	Tramadol, 150 mg	40		Tramadol, 100 mg	50
	APAP 650 & P100	39		ASA 650 & C60	48
	Codeine, 60 mg	41		Codeine, 60 mg	50
	Placebo	40		Placebo	49
TX	Tramadol, 75 mg	35	TI2	Tramadol, 75 mg	45
	Tramadol, 150 mg	36		Tramadol, 150 mg	45
	APAP 650 & P100	37		APAP 650 & P100	42
	Codeine, 60 mg	33		Codeine, 60 mg	47
	Placebo	36		Placebo	44
TY	Tramadol, 75 mg	31	TO	Tramadol, 100 mg	50
	Tramadol, 150 mg	28		Tramadol, 200 mg	50
	APAP 650 & P100	31		Codeine, 60 mg	47
	Codeine, 60 mg	30		Placebo	5
	Placebo	30			
TZA	Tramadol, 75 mg	39	TQ	Tramadol, 75 mg	50
	Tramadol, 150 mg	40		Tramadol, 150 mg	50
	APAP 650 & P100	38		APAP 650 & P100	49
	Codeine, 60 mg	38		Codeine, 60 mg	50
	Placebo	41		Placebo	51

ASA 650 & C60 = aspirin, 650 mg, plus codeine, 60 mg
APAP 650 & P100 = paracetamol, 650 mg, plus propoxyphene, 100 mg

TABLE 48 *Analgesic effectiveness of oral tramadol, codeine and combination analgesics in postoperative pain*

	Improved on active	Improved on control	RB (95% CI)	NNT (95% CI)
Dental				
Codeine, 60 mg	36/374	28/373	1.3 (0.8–2.1)	47.2 (16.3–∞)
Tramadol, 50 mg	41/246	13/225	2.9 (1.6–5.2)	9.2 (6.1–18.8)
Tramadol, 75 mg	16/95	6/95	2.7 (1.1–6.5)	9.5 (5.1–64.5)
Tramadol, 100 mg	89/300	22/278	3.8 (2.4–5.8)	4.6 (3.6–6.4)
Tramadol, 150 mg	29/95	6/95	4.8 (2.1–11.1)	4.1 (2.9–7.3)
Paracetamol, 650 mg, and propoxyphene, 100 mg	23/91	6/95	4.0 (1.7–9.4)	5.3 (3.4–11.4)
Aspirin, 650 mg, and codeine, 60 mg	52/235	13/225	3.8 (2.2–6.8)	6.3 (4.5–9.8)
Postsurgical				
Codeine, 60 mg	63/275	35/283	1.9 (1.3–2.7)	9.5 (6–23.4)
Tramadol, 50 mg	38/163	13/136	2.4 (1.4–4.4)	7.3 (4.6–17.9)
Tramadol, 75 mg	74/186	31/187	2.4 (1.7–3.5)	4.3 (3.1–7)
Tramadol, 100 mg	51/168	13/136	3.2 (1.8–5.6)	4.8 (3.4–8.2)
Tramadol, 150 mg	106/184	31/187	3.5 (2.5–4.9)	2.4 (2–3.1)
Paracetamol, 650 mg, and propoxyphene, 100 mg	91/225	34/227	2.7 (1.9–3.8)	3.9 (3–5.7)
Aspirin, 650 mg, and codeine, 60 mg	24/70	4/68	5.8 (2.1–15.9)	3.5 (2.5–6.3)
Combined				
Codeine, 60 mg	99/649	63/656	1.6 (1.2–2.1)	16.7 (11–48)
Tramadol, 50 mg	79/409	26/361	2.7 (1.8–4.1)	8.3 (6.0–13)
Tramadol, 75 mg	90/281	37/282	2.4 (1.7–3.5)	5.3 (3.9–8.2)
Tramadol, 100 mg	140/468	35/414	3.5 (2.5–5.0)	4.8 (3.8–6.1)
Tramadol, 150 mg	135/279	37/282	3.7 (2.7–5.1)	2.9 (2.4–3.6)
Paracetamol, 650 mg, and propoxyphene, 100 mg	114/316	40/322	2.9 (2.1–4.0)	4.2 (3.3–5.8)
Aspirin, 650 mg, and codeine, 60 mg	76/305	17/293	4.3 (2.6–7.1)	5.3 (4.1–7.4)

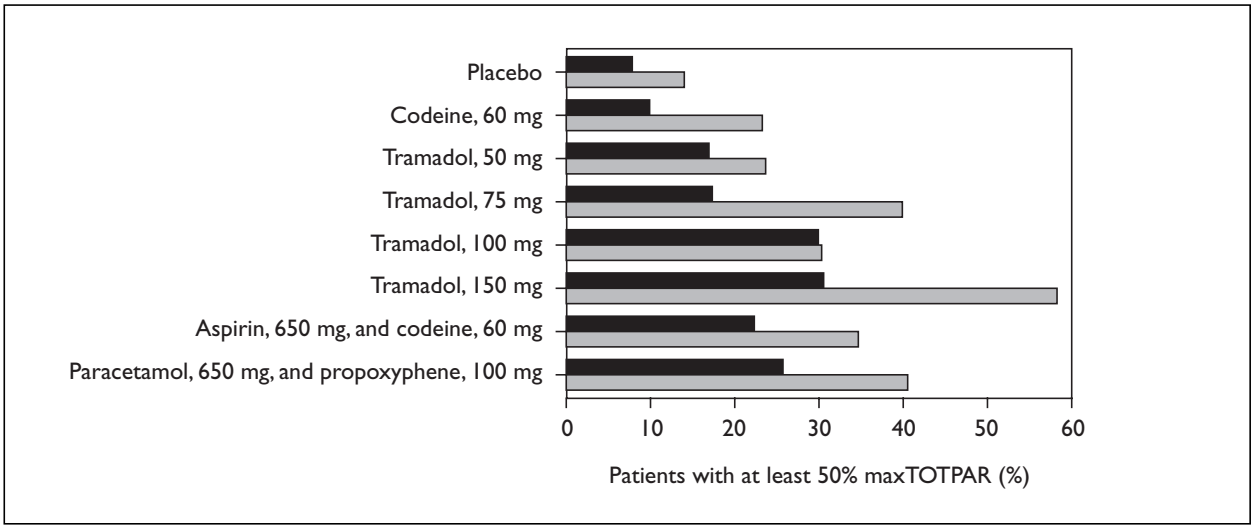


FIGURE 40 *Proportion of patients with at least 50% maxTOTPAR (■, dental; □, postoperative)*

This difference was not due to a greater proportion of postsurgical patients with more moderate than severe initial pain intensity. The ratio of moderate to severe initial (baseline) pain intensity was 931:663 (1.40:1) in postsurgical pain compared with 1294:505 (2.56:1) in dental pain (significantly less severe initial pain in the dental group, relative risk 0.81 (95% CI, 0.77–0.85)). Initial pain intensity stratification produced no consistent or significant differences in the proportion of patients with at least 50% maxTOTPAR, or in NNTs (data not shown). For postsurgical pain, 44 of

323 patients (13.6%) given placebo had at least 50% maxTOTPAR, compared with 28 of 373 patients (7.5%) for dental pain patients given placebo (relative benefit 1.8 (95% CI, 1.2–2.8).

Combined data

NNTs for dental and postsurgical patients combined are also shown in *Table 48*. Few data sets have sufficient information to allow calculation of analgesic efficacy for dental and postsurgical pain models separately. These numbers are those used for comparisons with other analgesic drugs from published reports without individual meta-analysis.

Choice of half pain relief

In order to test the effect of choices other than half pain relief, NNTs were calculated using dichotomous data for 20–80% maxTOTPAR (combined data). The results are shown in *Figure 41*.

NNTs for an effective drug, tramadol, 150 mg, were essentially the same, at about 2–3, over a wide range of decision points. Those for a slightly less effective analgesic (aspirin, 650 mg, plus codeine, 60 mg) rose slightly with pain relief cut-off values above 50% maxTOTPAR but for codeine, 60 mg, NNT values which started at about 10 for > 20% maxTOTPAR rose rapidly and were not significantly different from placebo by > 60% maxTOTPAR.

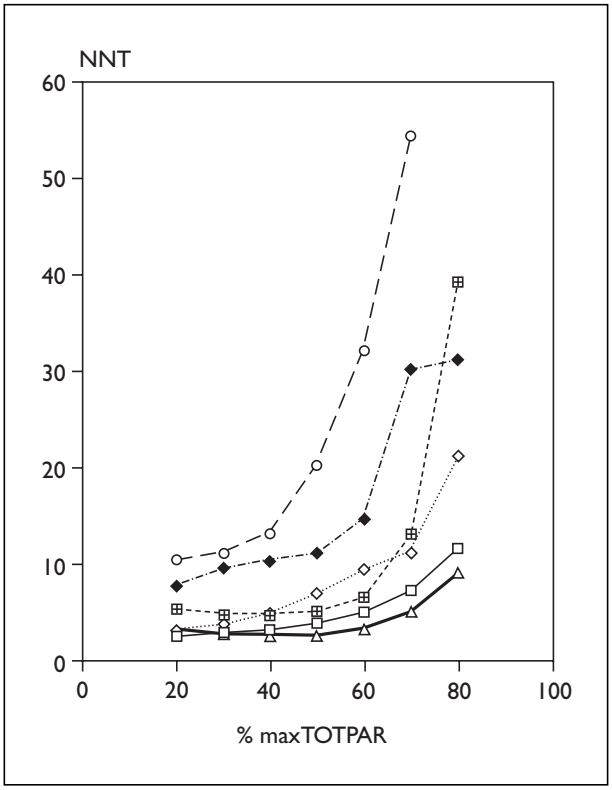


FIGURE 41 *Effect of % maxTOTPAR on NNT(——, paracetamol 650 mg/propoxyphene 100 mg; ....., aspirin 650 mg/codeine 60 mg; – – –, codeine 60 mg; ——, tramadol 150 mg; - - -, tramadol 100 mg; - · - · - ·, tramadol 50 mg)*

Adverse events

The incidence of the more common adverse events reported is shown in *Figure 42* for dental and postsurgical pain. Headache, vomiting, nausea, dizziness and somnolence were the most commonly reported adverse events, though predominantly of mild intensity.

For dental but not postsurgical pain, the adverse event incidence was generally sufficiently high to achieve a statistical difference from placebo for vomiting, nausea, dizziness and somnolence, but not headache. For tramadol there was a distinct dose response in dental pain, with higher doses producing greater incidence of adverse events; this trend was not present in postsurgical pain. NNHs can be calculated for adverse effects in dental patients because their incidence was sufficiently high. The clear dose response is shown in *Figure 43*.

Comment

Tramadol is an effective analgesic in postoperative pain. All doses of tramadol were statistically superior to placebo in both postsurgical and dental pain, and there was a significant dose response. Single oral doses of tramadol, 75 mg to 150 mg, had analgesic efficacy equivalent to combinations of paracetamol plus propoxyphene and aspirin plus codeine. Internal sensitivity was demonstrated by two comparator analgesics being statistically superior to placebo, and by the dose response for tramadol. The study methodology, randomised, double-blind trials, avoided known sources of major bias.

Search strategies identified a total of 20 randomised trials of oral tramadol in postoperative acute pain with standardised measurements of pain intensity and pain relief. Only two of these had been published in full[17,18] with one other in press.[22] Results of the other trials had been published in summary form only.[14] Meta-analytic tools so far developed have concentrated on patients achieving at least 50% pain relief as a single dichotomous measure of clinical effectiveness,[13] so the two studies which used pain intensity and not pain relief scales could not be included. Of the two studies excluded, one[18] could not distinguish oral tramadol from placebo after orthopaedic surgery; the other[22] found oral tramadol effective in dentoalveolar surgical pain.

Because the only two studies published in full[17,18] came to contrary conclusions about the efficacy of oral tramadol in postoperative pain, a controversy has arisen about its analgesic properties.[15,16] Examining all the available information (published and unpublished) has demonstrated clear analgesic efficacy; three standard analgesics were distinguished from placebo, as were tramadol, 50, 100 and 150 mg. Larger doses of tramadol produced more analgesia. This was performed in a single patient meta-analysis, a method which is claimed to be more conservative than aggregating mean data.[9]

Sunshine,[14] Cooper[20] and others have pointed out the variability that can occur in clinical trials of analgesics even in standard settings. This variability may have a number of causes. One is almost certain to be caused by the random play of chance in clinical trials where group sizes are of the order of 30 patients, although there may also be systematic causes. These may be apparent only in the systematic examination of large numbers of clinical trials with common endpoints.

The other consideration is that analgesic trial designs are explanatory.[12] They are designed to demonstrate that a particular compound is an analgesic in single doses in acute pain. They cannot in themselves determine the value of the intervention in clinical practice, although meta-analysis of such trials may be helpful in determining relative efficacy.[13]

This unique opportunity to use the individual patient data from 18 trials conducted to a common protocol allowed several questions to be addressed. The first, the original purpose of the studies, was to compare the efficacy and adverse effects of the novel analgesic with placebo and standard oral analgesics. The data also allowed the confirmation of the usefulness of at least 50% pain relief as an indicator of efficacy, comparison of the various analgesics in patients with either moderate or severe baseline pain and comparison of dental with postsurgical pain.

As a clinical outcome, 50% relief of pain has historical provenance over 40 years,[2] and is more readily clinically interpretable than the summary TOTPAR measure. For comparisons of analgesics, the question arises whether the 50% relief is a better cut-off than 20% or 80% relief. The best performing analgesic, tramadol, 150 mg, had NNT values of 2–3 across a range of cut-offs, from 30% to 60% maxTOTPAR *(Figure 43)*. The least analgesic drug, codeine, 60 mg, showed a rapid rise in NNT

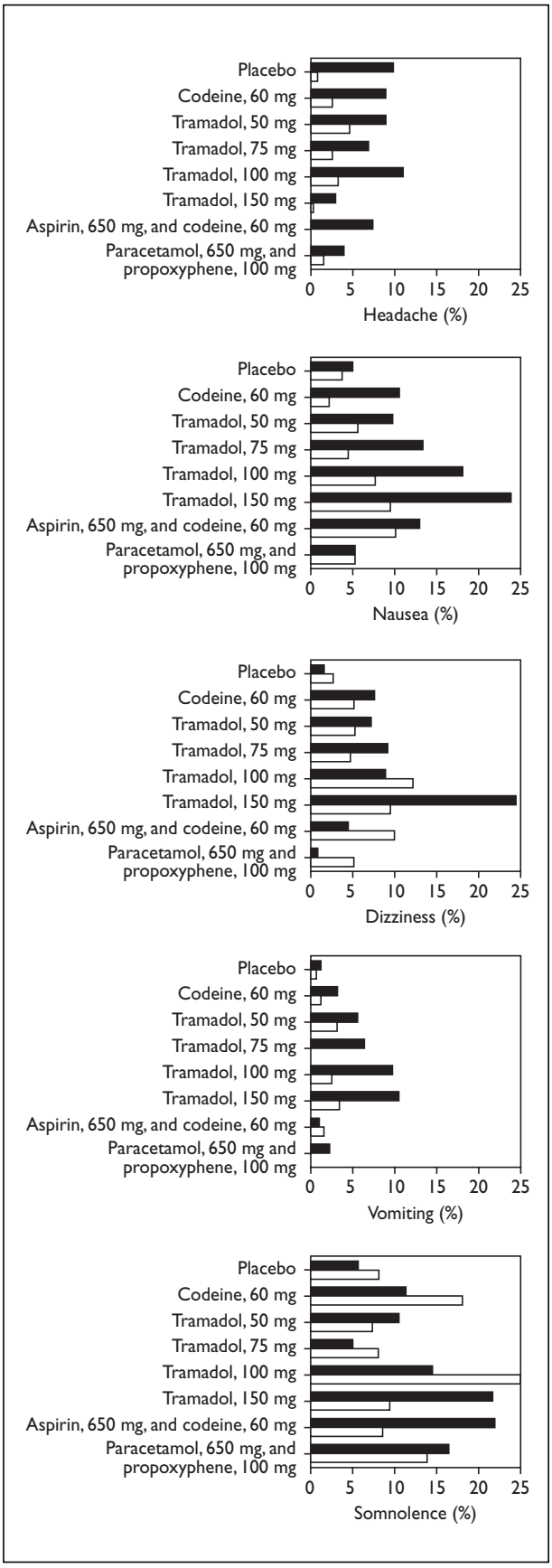


FIGURE 42 *Incidence of adverse events with oral tramadol – headache, nausea, vomiting, dizziness and somnolence (■, dental; □, postoperative)*

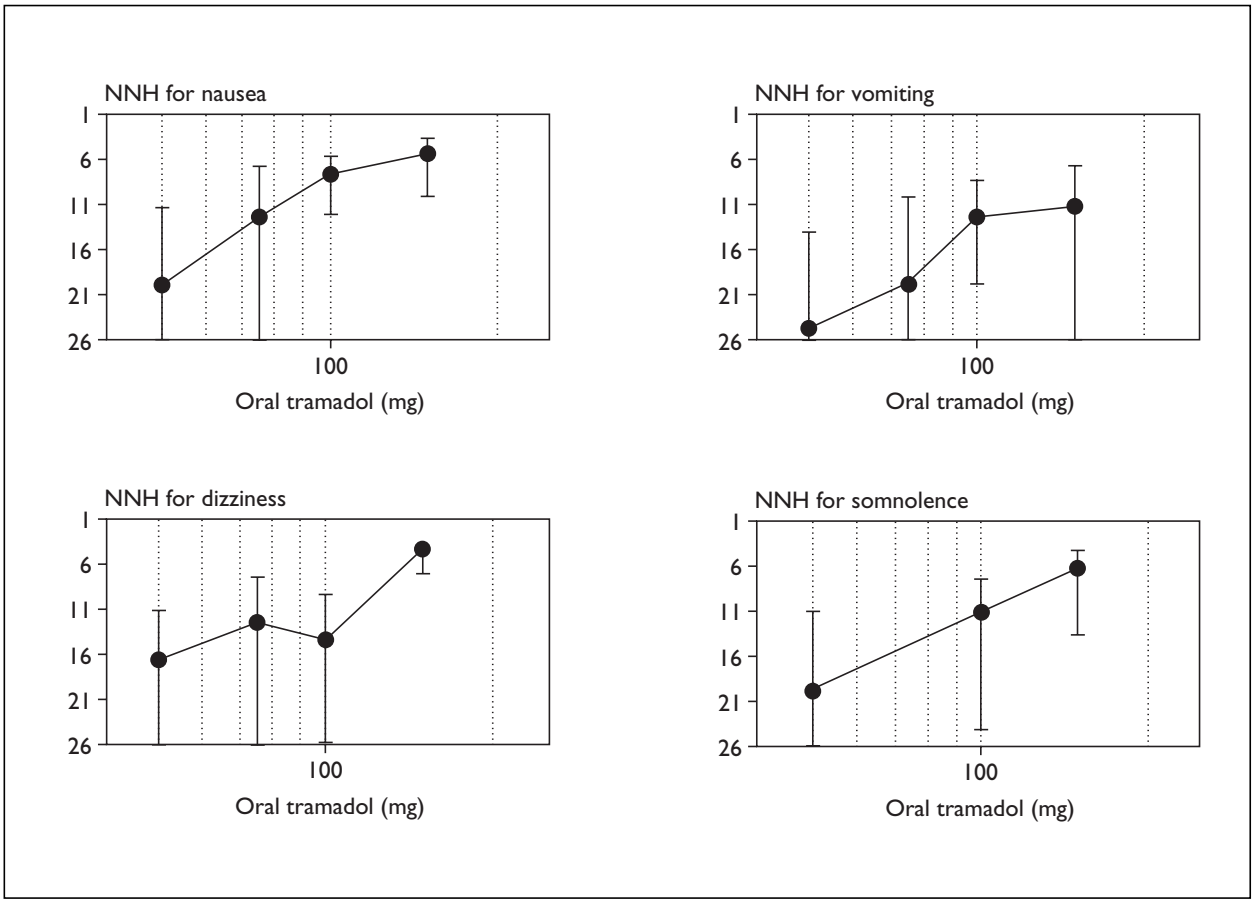


FIGURE 43 *NNH for adverse events in dental patients*

beyond 50% maxTOTPAR, and the aspirin–codeine combination showed a gradual rise in NNT with higher cut-offs. This provides some empirical support for the use of 50% as the cut-off. Not only does it have a clinically useful resonance but it also provides sensible discrimination between the best and worst analgesics.

That the NNT (the reciprocal of the absolute risk reduction, or risk difference) should be relatively unaffected by choice of cut-off is not unexpected. With placebo the proportion of patients achieving a particular level of pain relief falls quickly as % maxTOTPAR increases. For effective analgesics, this proportion falls slowly until high % maxTOTPAR levels are reached. The difference will remain largely unaltered over a wide range of % maxTOTPAR – generating stable NNTs.

The imposition of an arbitrary dichotomous outcome, at least 50% maxTOTPAR, on continuous data – a spectrum of response between no pain relief and complete pain relief – is justified because it allows analgesics to be compared across many different trials. However, it should not be overinterpreted; patients with less than 50% maxTOTPAR can also obtain useful pain relief; conversely those with at least 50% maxTOTPAR may have near maximal pain relief. The reality, though, is that multiple dosing is the norm in pain management, where adverse effects may drive practice as much as analgesia.

It has been suggested that differences might be seen between analgesics when tested on pain of initial moderate as opposed to severe intensity. This was not supported by these data. Stratification by initial pain intensity revealed no consistent or significant differences in the proportion of patients with at least 50% maxTOTPAR.

In these trials the analgesics were more effective in postsurgical pain than in dental pain *(Table 48),* producing lower NNTs despite there being significantly more patients with severe pain intensity at baseline in postsurgical pain. In postsurgical trials, significantly more patients given placebo (14%) had at least 50% pain relief than was the case with dental models (8%). These average figures for nine trials in each group are lower than those found by Cooper.[20]

Cooper's figures were derived from study mean TOTPAR. McQuay and colleagues[6] have pointed out that means are inadequate descriptors of asymmetrically distributed pain measurements, making comparison between estimates derived by single patient meta-analysis difficult. What Cooper's data did show was the great between-trial variability of placebo and active responses. This variability is not limited to acute pain studies, and is seen also in chronic pain studies,[23] as well as in studies with more objective outcomes like postoperative vomiting,[24] and in the response of infants to pulmonary surfactant.[25]

Despite analysing results on nearly 3500 patients, there were only 18 trials, nine each in dental and postsurgical pain. In order to make definitive statements about differences between pain models information from many more trials would need to be available. Cooper's 1991 analysis[20] had information from as many as 63 studies. The differences in analgesic efficacy and adverse events seen in this study support the view that dental and postoperative pain should be considered separately in meta-analytical comparisons of analgesic efficacy, at least when opioid analgesics or combinations with opioids are used.

Analyses of other analgesics in postsurgical and dental pain models are needed to allow comparisons of relative effectiveness to be made. This will not be easy, partly because few studies report data in ways which allow meta-analysis of the published reports, and partly because many patients are needed to obtain estimates with narrow CIs. Single patient meta-analysis is the most useful method of generating comparative information. However, it will involve much cooperation between clinical investigators and sponsoring pharmaceutical companies.

Authors of reports of trials of analgesics can aid future meta-analysis by including dichotomous outcomes as part of their analysis and report. This can easily be done as an addition to, not to the exclusion of, classical pain measures and analysis.[6]

References

1. Moore RA, McQuay HJ. Single-patient data meta-analysis of 3453 postoperative patients: oral tramadol versus placebo, codeine and combination analgesics. *Pain* 1997;**69**:287–94.

2. Denton J, Beecher H. New analgesics. *JAMA* 1949;**141**:1051–7.

3. Houde R, Wallenstein S, Beaver W. Clinical measurement of pain. In: de Stevens G, editor. Analgetics. New York: Academic Press, 1965; 75–122.

4. Jadad AR, Carroll D, Moore A, McQuay H. Developing a database of published reports of randomised clinical trials in pain research. *Pain* 1996;**66**:239–46.

5. Schulz KF, Chalmers I, Hayes RJ, Altman DG. Empirical evidence of bias: dimensions of methodological quality associated with estimates of treatment effects in controlled trials. *JAMA* 1995;**273**:408–12.

6. McQuay H, Carroll D, Moore A. Variation in the placebo effect in randomised controlled trials of analgesics: all is as blind as it seems. *Pain* 1996;**64**:331–5.

7. Cook RJ, Sackett DL. The number needed to treat: a clinically useful measure of treatment effect. *BMJ* 1995;**310**:452–4.

8. Chalmers I. Applying overviews and meta-analyses at the bedside. *J Clin Epidemiol* 1995;**48**:67–70.

9. Stewart LA, Parmar MKB. Meta-analysis of the literature or of individual patient data: is there a difference? *Lancet* 1993;**341**:418–22.

10. Harvey I, Peters TJ, Toth B. Meta-analysis [letter]. *Lancet* 1993;**341**:964.

11. Oxman AD, Clarke MJ, Stewart LA. From science to practice. Meta-analyses using individual patient data are needed. *JAMA* 1995;**274**:845–6.

12. Schwartz D, Lellouch J. Explanatory and pragmatic attitudes in therapeutic trials. *J Chronic Dis* 1967;**20**:637–48.

13. Moore A, McQuay H, Gavaghan D. Deriving dichotomous outcome measures from continuous data in randomised controlled trials of analgesics. *Pain* 1996;**66**:229–37.

14. Sunshine A. New clinical experience with tramadol. *Drugs* 1994;**47** suppl 1:8–18.

15. Vielvoye-Kerkmeer APE. Comments on Stubhaug, Grimstad and Breivik: *Pain* 1995;**62**:111–18 [letter]. *Pain* 1996;**64**:401–2.

16. Stubhaug A, Grimstad J, Breivik H. Reply to Vielvoye-Kerkmeer [letter]. *Pain* 1996;**64**:402.

17. Sunshine A, Olson NZ, Zighelboim I, DeCastro A, Minn FL. Analgesic oral efficacy of tramadol hydrochloride in postoperative pain. *Clin Pharmacol Ther* 1992;**51**:740–6.

18. Stubhaug A, Grimstad J, Breivik H. Lack of analgesic effect of 50 and 100 mg oral tramadol after orthopaedic surgery: a randomized, double-blind, placebo and standard active drug comparison. *Pain* 1995;**62**:111–18.

19. Jadad AR, Moore RA, Carroll D, Jenkinson C, Reynolds DJM, Gavaghan DJ, *et al.* Assessing the quality of reports of randomized clinical trials: is blinding necessary? *Control Clin Trials* 1996;**17**:1–12.

20. Cooper SA. Single-dose analgesic studies: the upside and downside of assay sensitivity. *Adv Pain Res Ther* 1991;**18**:117–24.

21. Gardner MJ, Altman DG. Statistics with confidence. London: British Medical Journal, 1989.

22. Collins M, Young I, Sweeney P, Fenn GC, Stratford ME, Wilson A, *et al.* The effect of tramadol on dento-alveolar surgical pain. *Br J Oral Maxillofac Surg;* in press.

23. McQuay HJ, Tramèr M, Nye BA, Carroll D, Wiffen PJ, Moore RA. A systematic review of antidepressants in neuropathic pain. *Pain* 1996;**68**:217–27.

24. Tramèr M, Moore A, McQuay H. Prevention of vomiting after paediatric strabismus surgery: a systematic review using the numbers-needed-to-treat method. *Br J Anaesth* 1995;**75**:556–61.

25. Soll JC, McQueen MC. Respiratory distress syndrome. In: Sinclair JC, Bracken ME, editors. Effective care of the newborn infant. Oxford: Oxford University Press, 1992;chapter 15, p. 333.

Chapter 16

Pain relief from intra-articular morphine after knee surgery

Summary

Reduction of postoperative pain by injecting opioid into the knee joint is believed to support the hypothesis of peripheral opioid receptor activation in inflammation. This systematic review of RCTs was designed to examine the evidence for this. Main outcomes were pain intensity and the use of supplementary analgesics. Efficacy of intra-articular bupivacaine against placebo was used as an index of internal sensitivity. Evidence of efficacy was sought in both early (0–6 hours after intra-articular injection) and late (6–24 hours) periods.

In all, 36 RCTs in knee surgery were found. Six had both a local anaesthetic control and placebo; four showed internal sensitivity and had at least one outcome showing efficacy of intra-articular morphine against placebo. Six studies compared intra-articular morphine with intravenous or intramuscular morphine or with intra-articular saline without a bupivacaine control. Four of the six studies showed greater efficacy for intra-articular morphine. There was no dose response evident. No quantitative analysis of pooled data was undertaken.

Intra-articular morphine may have some effect in reducing postoperative pain intensity and consumption of analgesics. These studies had significant problems in design, data collection, statistical analysis and reporting. Trials of better methodological quality are needed for a conclusive answer that intra-articular morphine is analgesic, and that any analgesia produced is clinically useful.

This chapter of the review has been published in full by Kalso and colleagues.[1]

Introduction

Intra-articular morphine has been used as a clinical test of the hypothesis that peripheral opioid receptors are activated in inflammation.[2] The judgement that exogenous opioids can provide effective postoperative analgesia has been taken as confirmation of the hypothesis.[2] Even though many studies and reviews have been published on this subject, consensus on whether intra-articular opioids offer clinically relevant pain relief is still lacking.

The issue of the sensitivity of analgesic measurement is particularly important. Over 40 years ago, Beecher[3] and Houde[4] described methods for measuring analgesic drugs which were sensitive and reproducible. Sensitive analgesic assays depended upon patients experiencing pain of moderate or severe intensity before test drug administration.

The aim of this systematic review, using the evidence from all RCTs, was to investigate the evidence for an analgesic effect of intra-articular morphine and to examine those features of trial methodology which influence judgement of experimental or clinical effectiveness.

Methods

RCTs of intra-articular opioids were sought systematically. A number of different search strategies in both MEDLINE (1966–May 1996), EMBASE and the Oxford Pain Relief Database (1950–94) were used, without language restriction. Search terms used included 'intra-articular', 'opiates', 'opioids' and 'morphine' and 'random*'.[5] Additional reports were identified from the reference lists of retrieved reports and from review articles. Unpublished reports, abstracts and reviews were not considered. Authors were not contacted for original data.

Reports considered

Reports were considered if they were randomised comparisons of intra-articular morphine with placebo (saline), or different doses of intra-articular morphine, or comparisons of intra-articular morphine with systemic (intravenous or intramuscular) morphine. Reports of direct comparisons of intra-articular morphine and local anaesthetic agents[6,7] were not considered. Reports of pethidine[8] were not considered because of potential confounding due to its local anaesthetic properties.

Each report which could possibly be described as an RCT was read independently by each of the authors and scored using a three-item quality scale.[9] The scale takes into account proper randomisation, double-blinding and reporting of drop-outs and withdrawals. Consensus was then achieved. Information was taken from each report on treatments and controls, types of surgery and anaesthesia, number of patients enrolled and analysed, study design, observation periods, outcome measures used for pain intensity and consumption of supplementary analgesics and adverse effects.

Validity and inclusion criteria

Pre-hoc validity criteria were number of patients per treatment group ≥ 10,[10] standardised methods of measuring pain intensity, and general anaesthesia. Spinal or epidural anaesthetics were not accepted, nor were infiltrations of local anaesthetic into the joints, because it was judged that low pain scores in the immediate postoperative period could render studies insensitive.

Two periods, early (up to 6 hours from the intra-articular injection) and late (6–24 hours) were defined for the evaluation of effectiveness.

Effectiveness was defined as a significant difference (as reported in the original trials) between the active and the control in pain intensity (early and late) or total consumption of rescue analgesics.

There was a *pre-hoc* agreement that an adequate description of internal sensitivity was a requirement for the demonstration of an analgesic action of intra-articular morphine. Such sensitivity would be derived (not necessarily exclusively) from a statistically significant difference between a known analgesic (intra-articular local anaesthetic) and placebo, from intra-articular morphine being different from placebo, or from a dose response for intra-articular morphine.

Quantitative analysis of morphine against placebo was planned.

Results

A total of 33 RCTs were found in 31 reports, studying nearly 1500 patients (about 900 of whom received morphine). All were in knee surgery. Two reports were in Danish, one in German and the rest in English.

The reasons for exclusion of studies were as follows:

- duplicate publications[11,12]
- the influence of tourniquet time on the efficacy of intra-articular morphine as the only outcome[13]
- number of patients per group less than ten[14,15]
- double-dummy technique not used for intramuscular administration[16] (study 2)
- control group not blinded[17]
- controls were intra-articular bupivacaine and unblinded lumbar plexus block only[18]
- spinal anaesthesia[19] (study 1)[20]
- epidural anaesthesia[21]
- operative intra-articular local anaesthetic[11,20,22,23] (study 2)[24]
- non-standardised anaesthesia (general anaesthesia, spinal or epidural anaesthesia)[25]
- inadequate standardisation of the timing of pain measurements.[26]

Details of the included studies are shown in *Table 49*. In all these trials morphine, 0.5–5 mg, was used as the intra-articular opioid. Controls used were bupivacaine, 0.25–0.5%, as the only intra-articular local anaesthetic, intra-articular saline or intravenous or intramuscular morphine, 1–2 mg. No quantitative analysis of pooled data was performed because results were presented as means, which are inadequate descriptors of assymmetrically distributed data.[27]

Morphine versus saline with bupivacaine as an index of internal sensitivity

Six studies compared intra-articular morphine with both bupivacaine and saline *(Table 49)*. One[28] was only analysed for an early effect (fewer than ten evaluable patients in the late period).

In two studies,[16,39] intra-articular bupivacaine could not be differentiated from intra-articular saline and the sensitivity of the analgesic assay was not proven. There was no difference between intra-articular morphine and saline in either.

Four trials[28–31] showed significantly lower VAS pain intensity scores with intra-articular bupivacaine compared with intra-articular saline during the early period (0–6 hours) and so had internal sensitivity.

All four sensitive studies reported early outcomes. Three of the four studies showed significantly lower early pain intensity scores after intra-articular morphine compared with intra-articular saline[28,29,31] *(Figure 44A)*.

In the late period, from 6 hours onwards, intra-articular morphine produced significantly lower pain intensity scores compared with placebo in all three evaluable sensitive studies[29–31] *(Figure 44B)*.

TABLE 49 *Pain relief from intra-articular morphine after knee surgery: included studies*

Study	Included or reason for exclusion	Quality (1–5)	Drugs and routes (number of patients)	Results: compared with placebo (saline)				
				VAS PI: early bupivacaine	**VAS PI: late bupivacaine**	**VAS PI: early morphine**	**VAS PI: late morphine**	**Analgesic consumption: 24-hour total**
Included studies with active control (bupivacaine) and placebo								
Björnsson, *et al.*, 1994[16] study 1	Included	2	i.a. morphine, 1 mg, in normal saline, 20 ml (21) i.a. normal saline, 20 ml, (19) i.a. 0.25% bupivacaine, 20 ml, (19) i.a. 0.25% bupivacaine, 20 ml, + morphine, 1 mg (19)	No difference between bupivacaine and normal saline at 0.5, 1, 1.5 or 2 hours.	No difference at 8, 24 or 48 hours.	No difference between morphine and normal saline at 0.5, 1, 1.5 or 2 hours.	No difference at 8, 24 or 48 hours.	No difference.
Haynes, *et al.*, 1994[28]	Included in early, excluded from late (inadequate number of patients per group)	3	i.a. normal saline, 40 ml (10) i.a. morphine, 1 mg, in 39 ml normal saline (10) i.a. 0.25% bupivacaine, 40 ml, + 1 in 200,000 adrenaline (10) i.a. 0.25% bupivacaine, 40 ml, + 1 in 200,000 adrenaline + morphine, 1 mg (10)	Bupivacaine better than normal saline: at 2 hours, $p = 0.01$; at 4 hours $p < 0.05$; at 6 hours, $p < 0.05$	n < 10	Morphine better than normal saline: at 2 hours, $p = 0.01$; at 4 hours $p < 0.05$; at 6 hours, $p < 0.05$.	n < 10	n < 10
Joshi, *et al.*, 1993a[29]	Included	2	i.a. morphine, 5 mg, in 25 ml (10) i.a. 0.25% bupivacaine, 25 ml (10) i.a. morphine, 5 mg, in 0.25% bupivacaine, 25 ml (10) i.a. normal saline, 25 ml (10)	Bupivacaine better than normal saline at 1, 2, 4 hours. No *p*-values given.	NSD (8 or 24 hours).	Morphine better than normal saline at 1, 2, 4 hours; statistics as for bupivacaine.	Morphine better than normal saline at 8, 24 hours; statistics as for bupivacaine.	Significance is not mentioned.
Karlsson, *et al.*, 1995[30]	Included	4	i.a. morphine, 1 mg, in 20 ml (10) i.a. 0.375% bupivacaine, 20 ml, (10) i.a. morphine, 1 mg, in 0.375% bupivacaine, 20 ml (10) i.a. normal saline, 20 ml (10)	Bupivacaine better than normal saline at 2, 4, 6 hour; no actual *p*-values given.	NSD (24 or 48 hours).	NSD (2, 4, 6 hours).	Morphine better than normal saline at 24, 48 hours; statistics as for bupivacaine.	Morphine significantly better than normal saline (0–24 hours and 24–48 hours). No p-value given.
McSwiney, *et al.*, 1993[31]	Included	2	i.a. morphine, 5 mg, in 0.25% bupivacaine, 12.5 ml, + normal saline, 12.5 ml (10) i.a. 0.25% bupivacaine, 25 ml (10) i.a. morphine, 5 mg, in normal saline, 25 ml (10) i.a. normal saline, 25 ml (10)	Bupivacaine better than normal saline at: 0.5, 1, 1.5, 2, 4 hours; no SEM/SD bars; no *p*-values.	Bupivacaine better than normal saline at 8, 12 hours; no SEM/SD bars; no *p*-values.	Morphine better than normal saline at 0.5, 1, 1.5, 2, 4 hours; no SEM/SD bars; no *p*-values.	Morphine better than normal saline at 8, 12, 24 hours; no SEM/SD bars; no *p*-values.	Morphine and bupivacaine significantly better ($p < 0.05$) than normal saline.
Included studies with placebo but no local anaesthetic as active control								
Joshi, *et al.*, 1992[32]	Included	2	i.a. morphine, 5 mg, in 25 ml normal saline (10) i.a. normal saline, 25 ml (10)			i.a. morphine better than i.a. normal saline at 0, 0.5, 1, 1.5, 2 and 4 hours; $p < 0.05$	i.a. morphine better than i.a. normal saline at 8 and 12 hours; $p < 0.05$.	i.a. morphine better than i.a. normal saline; $p < 0.05$.
Joshi, *et al.*, 1993[40]	Included	2	i.a. morphine, 5 mg, in 25 ml normal saline (10) i.a. normal saline, 25 ml (10)			NSD (1, 2 or 4 hours).	NSD (8 or 24 hours).	i.a. morphine better than i.a. normal saline, $p < 0.01$.
Lyons, *et al.*, 1995[41]	Pethidine arm not considered	3	i.a. morphine, 5 mg, in 25 ml normal saline (20) i.a. normal saline, 25 ml (20) i.a. pethidine, 50 mg, in 25 mg normal saline (20)			i.a. morphine better than i.a. normal saline at 0, 0.5, 1, 2 and 4 hours; $p < 0.01$.	i.a. morphine better than i.a. normal saline at 8, 12 and 24 hours; $p < 0.01$.	
Included studies; cross-route morphine								
Dierking, *et al.*, 1994[33]	Included	4	i.a. morphine, 2 mg, in normal saline, 40 ml, + i.m. normal saline, 1 ml (18) i.a. normal saline, 40 ml, + i.m. morphine, 2 mg (15)			NSD (1, 2, 4, 6 hours).	Not evaluated.	Not evaluated.

continued

TABLE 49 contd *Pain relief from intra-articular morphine after knee surgery: included studies*

Study	Included or reason for exclusion	Quality (1–5)	Drugs and routes (number of patients)	Results: compared with placebo (saline)				
				VAS PI: early bupivacaine	VAS PI: late bupivacaine	VAS PI: early morphine	VAS PI: late morphine	Analgesic consumption: 24-hour total
Included studies; cross-route morphine contd								
Hege-Scheuing, et al., 1995[34]	Included	3	i.a. morphine, 1mg, in 10 ml + i.v. normal saline, 10 ml (29) i.v. morphine, 1 mg, in 10 ml + i.a. normal saline, 10 ml (30)			NSD (1, 2, 3, 4, 6 hours).	NSD (8 or 24 hours).	No difference.
Stein, *et al.*, 1991[35]	Included (low dose morphine excluded), inadequate number of patients	2	i.a. morphine, 1 mg, in normal saline, 40 ml, + i.v. normal saline, 1 ml (18) i.a. normal saline, 40 ml, + i.v. morphine, 1 mg (15) i.a. morphine, 0.5 mg, in normal saline, 40 ml, + i.v. normal saline, 1 ml (10) i.a. morphine, 1 mg, + naloxone, 0.1 mg, in normal saline, 40 ml, + i.v. normal saline, 1 ml (9)			i.a. morphine better than i.v. morphine at 3, 4, 6 hours; $p < 0.05$.	NSD (24 hours).	i.a. morphine significantly better than i.v. morphine.
Included studies: morphine dose response								
Allen, *et al.*, 1993[36]	Included	5	i.a. 0.25% bupivacaine, 30 ml (30) i.a. morphine, 1 mg, in normal saline, 30 ml (30) i.a. morphine, 2 mg, in normal saline, 30 ml (30) i.a. morphine, 1 mg, in 0.25% bupivacaine, 30 ml (30)					
Heine, *et al.*, 1994[37]	Included	2	i.a. 0.5% bupivacaine, 20 ml (11) i.a. morphine, 1 mg, in 0.5% bupivacaine, 20 ml (10) i.a. morphine, 3 mg, in 0.5% bupivacaine, 20 ml (10)					
Juelsgaard, *et al.*, 1993[23]	Included	3	i.a. morphine, 2 mg, in 5 ml (25) i.a. morphine, 4 mg in 5 ml (25)					
Laurent, *et al.*, 1994[38]	Included	4	i.a. morphine, 5 mg, in 0.25% bupivacaine, 40 ml (20) i.a. morphine, 2 mg, in 0.25% bupivacaine, 40 ml (20) i.a. 0.25% bupivacaine, 40 ml (18)					
Included studies: not analysed								
Khoury, *et al.*, 1992[6]	Included (early), excluded (late) inadequate number of patients per group	3	i.a. morphine, 1 mg, in 20 ml (11) i.a. 0.25% bupivacaine, 20 ml (11) i.a. morphine, 1 mg, + 0.25% bupivacaine, 20 ml (11)					
VanNess, *et al.*, 1994[7]	Included	1	i.a. 0.25% bupivacaine, 30 ml, + 1 in 200 k adrenaline (41) i.a. morphine, 2 mg (2 ml), + normal saline, 28 ml (40)					

Total consumption of supplementary analgesics over 24 hours was significantly lower after intra-articular morphine compared with saline in the two sensitive studies in which it was analysed.[30,31]

Morphine versus saline, no active (bupivacaine) control

Three studies compared only intra-articular morphine with intra-articular saline.[32,40,41]

In the early period, morphine VAS pain intensity scores were significantly lower in two of the three studies which reported early outcomes[32,41] *(Figure 44C)*.

In the late period, the same two studies[32,41] indicated that intra-articular morphine produced significantly lower pain intensity scores compared with saline. Two of the three studies[32,40] had a

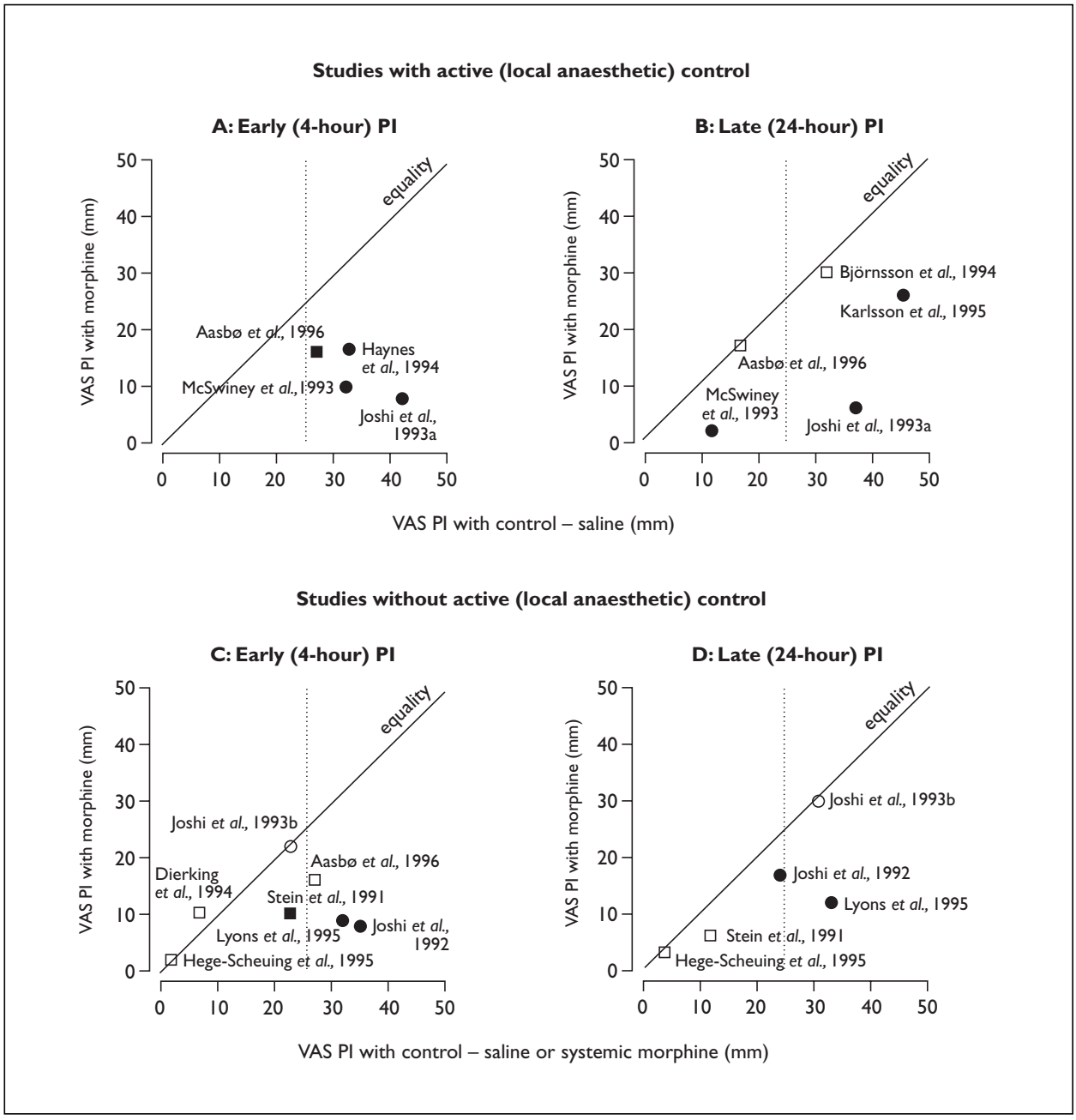


FIGURE 44 *L'Abbé plots for intra-articular morphine trials*

significantly lower total consumption of analgesics over 24 hours after morphine.

Morphine versus systemic morphine control

Three studies compared intra-articular with intravenous or intramuscular morphine.[33–35]

In the early period, one study[35] showed greater efficacy for intra-articular morphine compared with intravenous morphine, 1 mg *(Figure 44C)*.

In the late period, no study indicated that intra-articular morphine had statistically lower pain intensity scores, although in one study[33] there were no evaluations beyond 6 hours *(Figure 44D)*. Lower total consumption of analgesics over 24 hours was found in only one study.[35]

Combination of morphine plus bupivacaine versus saline

All four sensitive studies which compared intra-articular morphine with both saline and bupivacaine also included a group with a combination of intra-articular morphine plus bupivacaine. All the studies which were sensitive to bupivacaine alone and showed a positive effect for morphine

also showed a significant effect for the combination compared with placebo, both early and late.[28–31] The two studies which were insensitive for bupivacaine and morphine showed no efficacy for the combination.[16,39]

Dose response

Two studies addressed the question of a dose response with intra-articular morphine alone.[35,36] In one study,[35] intra-articular morphine, 1 mg, could be differentiated from control but not 0.5 mg; the researchers could not differentiate 0.5 mg from 1 mg of morphine. The other study[36] demonstrated a reversed dose response between 1 and 2 mg. Neither study had evidence of internal sensitivity.

In two studies different doses of morphine in combination with a standard dose of bupivacaine were compared.[37,38] No dose response was detected between either 1 and 3 mg or 2 and 5 mg of morphine.

Adverse effects

No adverse effects that could have been attributed to the intra-articular treatment were reported.

Comment

These reports of the use of intra-articular morphine emphasise the importance of considering potential bias and issues of validity in clinical studies before interpreting results.

Bias

It is now well recognised that studies which are either not randomised or randomised without concealment of treatment allocation, or which are not adequately blinded, result in an overestimation of the effect of treatment.[42] Method and concealment of randomisation, double-blinding and withdrawals and dropouts were inadequately described in all these studies. The method of randomisation was explicit in three studies.[30,36,38] In many studies, it was unclear who was blinded.

Design and validity

Classic analgesic trial design includes both active and placebo controls. The reason is to ensure that if no difference is found between test analgesic and placebo, the correct interpretation of a negative result can be made if the standard (active control) analgesic gives a significant difference from placebo. This is particularly important when pain is of only mild to moderate intensity. The mean pain intensities after placebo were, with one exception, less than 50% of the maximum possible,[30] both early and late, and frequently below 25% of maximum *(Figure 44)*. If there is no pain, reduction in pain intensity cannot be measured. The reduced sensitivity of analgesic studies with low pain intensity has been evaluated.[43]

For this reason a hierarchy of evidence was chosen. The highest rank was when active (bupivacaine) control was used as well as placebo, and analgesic efficacy of intra-articular morphine was interpreted only when intra-articular bupivacaine was more effective than placebo (i.e. established internal sensitivity). Intra-articular bupivacaine is known to provide reliable analgesia of predictable duration following knee surgery,[44,45] and it was therefore a valid active control.

Outcome measures

The special feature of these studies was that necessarily the intervention was made before the patient had pain, analogous to pre-emptive studies.[46] The VAS pain intensity levels were low for several reasons. Diagnostic arthroscopies were included in the primary studies, and opioids and NSAIDs were given both pre- and perioperatively. Diagnostic arthroscopies may not cause enough postoperative pain to be sufficiently sensitive for an analgesic assay. *Figure 44* indicates that studies which were sensitive generally had VAS pain intensity levels above 30% of the maximum possible in the control group in the early period (and most had high values in the late period also). We excluded studies which used spinal or epidural anaesthesia or infiltration of the knee joint with high doses of local anaesthetic because these measures further reduce postoperative pain and, hence, sensitivity.

VAS pain intensity was usually measured at rest; sensitivity might have been increased by assessing it on movement. Arthroscopic surgery is usually performed as day-case surgery. Sensitivity might have been increased by following the patients in hospital for a longer period. Patients were instructed in the use of VAS PI before anaesthesia in only a minority of studies. Most patients were sent home with a questionnaire within 2–6 hours of the end of surgery. Few studies mentioned whether the assessment was performed by a trained, or even the same, observer. All these issues should be addressed in study design. Sensitivity of the analgesic assay is crucial.

Consumption of supplementary analgesics within the first 24 hours after surgery was the second commonest outcome measure but usually was not standardised. Other indicators of pain and pain

relief, such as time to first analgesic, time to weight bearing, time to discharge, were also used, but in only a minority of studies. VAS pain intensity and the total consumption of supplementary analgesics were therefore used as primary outcome measures in the analysis.

Early and late periods

Analysis by early and late periods was used for several reasons. For the first 2–6 hours patients were still in hospital where VAS pain intensity measurements were made by researchers or (trained) nurses at predetermined intervals. Secondly, the effect of intra-articular bupivacaine, the index of internal trial sensitivity, should have been most pronounced over this time. Thirdly, any systemic effect of morphine should have been more obvious during this period rather than later. The late period was considered to be important as several studies suggested a prolonged effect of intra-articular morphine. Most studies provided information on VAS pain intensity values at 24 hours and consumption of supplementary analgesics was reported as a total amount taken over 24 hours.

No biological reason for suspecting a late rather than an early effect was apparent in the original study on intra-articular morphine.[35] This indicated that intra-articular morphine, 1 mg, provided significantly better analgesia after knee surgery than the same dose given intravenously at 3, 4 or 6 hours. No difference was found between VAS pain intensity values at 24 hours, although the total consumption of supplementary analgesics during the 24-hour period was significantly less after intra-articular morphine.

Studies with both active and placebo controls

Only six studies included groups receiving saline, bupivacaine and morphine. Four of the studies were judged sensitive as defined by significant analgesic effect of bupivacaine compared with saline. All four studies demonstrated significant analgesic effect of intra-articular morphine compared with placebo at both early and late times *(Figure 44)*. This provides some evidence for a prolonged biological effect of morphine in the knee joint. The two negative studies[16,39] failed the sensitivity test.

Studies with no active control

Three studies of morphine against saline showed an analgesic effect, two in the early period and all three in the late period. Comparisons of intra-articular with intravenous or intramuscular morphine were less compelling; only one of the three studies[35] showed a significant effect, both early and late. These results again provide some evidence for an analgesic effect of morphine in the knee joint, while raising the issue of whether this is a systemic as opposed to a local effect.

Dose–response studies

No dose response was detectable in any study, over a dose range of 0.5–5 mg. The minimum dose tested (0.5 mg) did not show analgesic efficacy[35] whereas a dose of 1 mg did.[28,30,35] No greater effect was found using morphine doses of 2 mg compared with 1 mg.[36] In combination with local anaesthetic, morphine doses of 3 mg compared with 1 mg[37] and 5 mg compared with 2 mg[38] showed no increased efficacy. None of these studies had proven internal sensitivity. Failure to demonstrate dose response may therefore have been due to lack of sensitivity in the methods.

However, the lowest effective dose of morphine used, 1 mg, would, in a 20 ml injection, be equivalent to a concentration of about 200 μmol per litre (50 μg/ml). Typical blood or tissue levels after systemic injections of analgesic doses of morphine are found at concentrations of nanomols per litre, at least 1000 times lower.[47] The very high concentrations of morphine in the knee joint would be expected to saturate any opioid receptors present. If morphine is acting on local opioid receptors, then the minimal effective dose may well be much less than 1 mg. Failure to demonstrate a dose response might then be because the doses tested were at the top end of the dose–response curve. Late efficacy might be a consequence of residual high morphine concentrations.

Is intra-articular morphine effective?

Taken together, these results render some support for the hypothesis that intra-articular morphine provides pain relief after knee surgery.[2] Using a simple 'vote-counting' approach on *Figure 44,* the points from the majority of the trials fall in the lower right quadrant, indicating greater efficacy with morphine than control. Convincing evidence for an early effect is lacking. There was more consistent evidence for a prolonged analgesic effect, mostly a single estimate of pain intensity at 24 hours or consumption of analgesic. These are weak measures.

The problem is that this evidence rests on four trials which fulfilled the sensitivity requirements but which had only ten patients per treatment group, and two others which were methodologically weak but did distinguish morphine from saline. Against these studies stands the failure to demonstrate a dose response for intra-articular morphine.

Overall, the evidence is not compelling. The lessons for future studies are obvious, but the current agenda is one of research rather than clinical utility.

References

1. Kalso E, Tramèr M, Carroll D, McQuay H, Moore RA. Pain relief from intra-articular morphine after knee surgery: a qualitative systematic review. *Pain* 1997;**71**:642–51.
2. Stein C. The control of pain in peripheral tissue by opioids. *N Engl J Med* 1995;**332**:1685–90.
3. Beecher HK. The powerful placebo. *JAMA* 1955;**159**:1602–6.
4. Houde R. On appraising pain and analgesic drugs. In: Keele C, Smith R, editors. The assessment of pain in man and animals. Edinburgh: Livingstone, 1962;202–12.
5. Jadad AR, McQuay HJ. A high-yield strategy to identify randomized controlled trials for systematic reviews. *Online J Curr Clin Trials* [serial online] 1993; Doc No 33:3973 words; 39 paragraphs. 5 tables.
6. Khoury GF, Chen ACN, Garland DE, Stein C. Inta-articular morphine, bupivacaine, and morphine/bupivacaine for pain control after knee videoarthroscopy. *Anesthesiology* 1992;**77**:263–6.
7. VanNess SA, Gittins ME. Comparison of intra-articular morphine and bupivacaine following knee arthroscopy. *Orthop Rev* 1994;**23**:743–7.
8. Ekblom A, Westman L, Söderlund A, Valentin A, Eriksson E. Is intra-articular pethidine an alternative to local anaesthetics in arthroscopy? A double-blind study comparing prilocaine with pethidine. *Knee Surg Sports Traumatol Arthrosc* 1993;**1**:189–94.
9. Jadad AR, Moore RA, Carroll D, Jenkinson C, Reynolds DJM, Gavaghan DJ, *et al.* Assessing the quality of reports of randomized clinical trials: is blinding necessary? *Control Clin Trials* 1996;**17**:1–12.
10. L'Abbé KA, Detsky AS, O'Rourke K. Meta-analysis in clinical research. *Ann Intern Med* 1987;**107**:224–33.
11. Dalsgaard J, Felsby S, Juelsgaard P, Frokjaer J. Analgesi af lavdosis intraartikulaer morfin efter ambulant knaeartroskopi [Analgesic effect of low-dose intra-articular morphine after ambulatory knee arthroscopy]. *Ugeskr Laeger* 1993;**155**:4166–9.
12. Lehrberger K, Stein C, Hassan A, Yassouridis A. Opioids as novel intra-articular agents for analgesia following arthroscopic knee surgery. *Knee Surg Sports Traumatol Arthrosc* 1994;**2**:174–5.
13. Klinken C. Effects of tourniquet time in knee arthroscopy patients receiving intra-articular morphine combined with bupivacaine. *CRNA* 1995;**6**:37–42.
14. Boden BP, Fassler S, Cooper S, Marchetto PA, Moyer RA. Analgesic effect of intra-articular morphine, bupivacaine, and morphine/bupivacaine after arthroscopic knee surgery. *Arthroscopy* 1994;**10**:104–7.
15. Joshi GP, McCarroll SM, Brady OH, Hurson BJ, Walsh G. Intra-articular morphine for pain relief after anterior cruciate ligament repair. *Br J Anaesth* 1993;**70**:97–8.
16. Björnsson A, Gupta A, Vegfors M, Lennmarken C, Sjöberg F. Intra-articular morphine for post-operative analgesia following knee arthroscopy. *Reg Anesth* 1994;**19**:104–8.
17. Chan ST. Intra-articular morphine and bupivacaine for pain relief after therapeutic arthroscopic knee surgery. *Singapore Med J* 1995;**36**:35–7.
18. De Andrés J, Bellver J, Barrera L, Febre E, Bolinches R. A comparative study of analgesia after knee surgery with intra-articular bupivacaine, intra-articular morphine, and lumbar plexus block. *Anesth Analg* 1993;**77**:727–30.
19. Ho ST, Wang JJ, Liaw WJ, Wong CS, Cherng CH. Analgesic effect of intra-articular morphine after arthroscopic knee surgery in Chinese patients. *Acta Anaesthesiol Sin* 1995;**33**:79–84.
20. Niemi L, Pitkänen M, Tuominen M, Björkenheim JM, Rosenberg PH. Intra-articular morphine for pain relief after knee arthroscopy performed under regional anaesthesia. *Acta Anaesthesiol Scand* 1994;**38**:402–5.
21. Raja SN, Dickstein RE, Johnson CA. Comparison of postoperative analgesic effects of intra-articular bupivacaine and morphine following arthroscopic knee surgery. *Anesthesiology* 1992;**77**:1143–7.
22. Jaureguito JW, Wilcox JF, Cohn SJ, Thisted RA, Reider B. A comparison of intra-articular morphine and bupivacaine for pain control after outpatient knee arthroscopy. A prospective, randomized, double-blinded study. *Am J Sports Med* 1995;**23**:350–3.
23. Juelsgaard P, Dalsgaard J, Felsby S, Frokjaer J. Analgetisk effekt af to doser intraartikulaer morfin efter ambulant knaeartroskopi. Et randomiseret, prospektivt, dobbeltblindt studie. *Ugeskr Laeger* 1993;**155**:4169–72.
24. Reuben SS, Connelly NR. Postarthroscopic meniscus repair analgesia with intra-articular ketorolac or morphine. *Anesth Analg* 1996;**82**:1036–9.
25. Heard SO, Edwards WT, Ferrari D, Hanna D, Wong PD, Liland A, *et al.* Analgesic effect of intra-articular bupivacaine or morphine after arthroscopic knee surgery: a randomized, prospective, double-blind study. *Anesth Analg* 1992;**74**:822–6.
26. Ruwe PA, Klein I, Shields CL. The effect of intra-articular injection of morphine and bupivacaine on postarthroscopic pain control. *Am J Sports Med* 1995;**23**:59–64.

27. McQuay H, Carroll D, Moore A. Variation in the placebo effect in randomised controlled trials of analgesics: all is as blind as it seems. *Pain* 1996;**64**:331–5.

28. Haynes TK, Appadurai IR, Power I, Rosen M, Grant A. Intra-articular morphine and bupivacaine analgesia after arthroscopic knee surgery. *Anaesthesia* 1994;**49**:54–6.

29. Joshi GP, McCarroll SM, O'Brien TM, Lenane P. Intra-articular analgesia following knee arthroscopy. *Anesth Analg* 1993;**76**:333–6.

30. Karlsson J, Rydgren B, Eriksson B, Jarvholm U, Lundin O, Sward L, *et al.* Postoperative analgesic effects of intra-articular bupivacaine and morphine after arthroscopic cruciate ligament surgery. *Knee Surg Sports Traumatol Arthrosc* 1995;**3**:55–9.

31. McSwiney MM, Joshi GP, Kenny P, McCarroll SM. Analgesia following arthroscopic knee surgery. A controlled study of intra-articular morphine, bupivacaine or both combined. *Anaesth Intensive Care* 1993;**21**:201–3.

32. Joshi GP, McCarroll SM, Cooney CM, Blunnie WP, O'Brien TM, Lawrence AJ. Intra-articular morphine for pain relief after knee arthroscopy. *J Bone Joint Surg Br* 1992;**74**:749–51.

33. Dierking GW, Østergaard MD, Dissing CK, Kristensen JE, Dahl JB. Analgesic effect of intra-articular morphine after arthroscopic meniscectomy. *Anaesthesia* 1994;**49**:627–9.

34. Hege-Scheuing G, Michaelsen K, Bühler A, Kustermann J, Seeling W. Analgesie durch intraartikuläres Morphin nach Kniegelenks-arthroskopien? *Anaesthesist* 1995;**44**:351–8.

35. Stein C, Comisel K, Haimerl E, Yassouridis A, Lehrberger K, Herz A, *et al.* Analgesic effect of intra-articular morphine after arthroscopic surgery. *N Engl J Med* 1991;**325**:1123–6.

36. Allen GC, St. Amand MA, Lui ACP, Johnson DH, Lindsay MP. Postarthroscopy analgesia with intra-articular bupivacaine/morphine. *Anesthesiology* 1993;**79**:475–80.

37. Heine MF, Tillet ED, Tsueda K, Loyd GE, Schroeder JA, Vogel RL, *et al.* Intra-articular morphine after arthroscopic knee operation. *Br J Anaesth* 1994;**73**:413–15.

38. Laurent SC, Nolan JP, Pozo JL, Jones CJ. Addition of morphine to intra-articular bupivacaine does not improve analgesia after day-case arthroscopy. *Br J Anaesth* 1994;**72**:170–3.

39. Aasbø V, Raeder JC, Grøgaard B, Røise O. No additional analgesic effect of intra-articular morphine or bupivacaine compared with placebo after elective knee arthroscopy. *Acta Anaesthesiol Scand* 1996;**40**:585–8.

40. Joshi GP, McCarroll SM, McSwiney M, O'Rourke P, Hurson BJ. Effects of intra-articular morphine on analgesic requirements after anterior cruciate ligament repair. *Reg Anesth* 1993;**18**:254–7.

41. Lyons B, Lohan D, Flynn CG, Joshi GP, O'Brien TM, McCarroll M. Intra-articular analgesia for arthroscopic meniscectomy. *Br J Anaesth* 1995;**75**:552–5.

42. Schulz KF, Chalmers I, Hayes RJ, Altman DG. Empirical evidence of bias: dimensions of methodological quality associated with estimates of treatment effects in controlled trials. *JAMA* 1995;**273**:408–12.

43. Stubhaug A, Grimstad J, Breivik H. Lack of analgesic effect of 50 and 100 mg oral tramadol after orthopaedic surgery: a randomized, double-blind, placebo and standard active drug comparison. *Pain* 1995;**62**:111–18.

44. Chirwa SS, MacLeod BA, Day B. Intra-articular bupivacaine (Marcaine) after arthroscopic meniscectomy: a randomized double blind controlled study [see comments]. *Arthroscopy* 1989;**5**:33–5.

45. Kaeding CC, Hill JA, Katz J, Benson L. Bupivacaine use after knee arthroscopy: pharmacokinetics and pain control study. *Arthroscopy* 1990;**6**:33–9.

46. McQuay HJ. Pre-emptive analgesia: a systematic review of clinical studies. *Ann Med* 1995;**27**:249–56.

47. Moore RA, Hand CW, McQuay HJ. Opioid metabolism and excretion. In: Budd K, editor. Update in Opioids. Clinical anaesthesiology. London: Baillère Tindall, 1987:829–58.

Chapter 17

Analgesic efficacy of peripheral opioids

Summary

Anaesthetists, using basic scientific concepts of peripheral opioid activity, try to improve regional anaesthesia and postoperative analgesia by injecting opioids, with or without local anaesthetic, close to nerve trunks or nerve endings. The aim of this systematic review was to test the evidence that peripherally applied opioids (all except intra-articular) have an analgesic effect outside the knee joint.

A systematic search, 1966–96, was carried out for published reports of RCTs which compared the efficacy of peripheral opioids with placebo, local anaesthetic, or systemic opioids in acute pain. Reports of pethidine or intra-articular opioids were not included. Data on intraoperative efficacy (onset, quality, duration of sensory block), and postoperative efficacy (pain intensity, analgesic consumption) were extracted. Statistical significance as indicated in the original reports and clinical relevance of differences between opioids and controls were taken into account to estimate qualitatively overall efficacy.

Data for 952 patients in 26 trials were analysed. The opioids used were morphine (16 trials), fentanyl (8), alfentanil (1), buprenorphine (1), and butorphanol (1). Two from four experimental pain trials reported a statistically significant difference in favour of the opioid. In 22 clinical trials, efficacy of opioid injections into the brachial plexus (10 trials), Bier's block (4), perineural (3), and other sites (5) was tested.

In five of ten clinical trials measuring intraoperative efficacy, statistically significant efficacy with opioids compared with control was reported; none were judged clinically relevant. In five of 17 clinical trials measuring postoperative efficacy, a significant difference in favour of the opioid was reported; none were judged clinical relevant. Trials of lower quality were more likely to report increased efficacy with opioids. Adverse events related to route of administration were not reported.

These trials provide no evidence for clinically relevant peripheral analgesic efficacy of opioids in acute pain.

This chapter of the review has been published in full by Picard and colleagues.[1]

Introduction

For over 10 years anaesthetists have been trying to improve efficacy of regional anaesthesia and postoperative analgesia by injecting opioids close to the nerve trunks or the nerve endings. The biological basis for this approach is the presence of opioid receptors and their endogenous ligands in the peripheral nervous system, and their effect on modulation of inflammatory pain.[2]

There are several distinct clinical approaches to this topic. First, do opioids, when injected in combination with local anaesthetics, improve the quality and duration of a sensory block? This could lead to improved surgical conditions. Second, does this method allow the dose of the local anaesthetic to be reduced? This would minimise the risk of systemic toxicity of local anaesthetics. Third, do opioids, when applied alone in peripheral sites, decrease postoperative pain intensity and analgesic requirements? This is a purer test of the biological question of whether opioids have analgesic effects peripherally.

The aim of this systematic review was to test the evidence that peripheral opioids (all except intra-articular) improve the quality of either intraoperative regional anaesthesia or postoperative analgesia

Methods

Full published reports of RCTs of peripheral opioids were sought systematically. A number of different search strategies in MEDLINE (1966–September 1996), EMBASE (1981–96) and the Oxford Pain Relief Database (1950–94) were used, without language restriction. Additional reports were identified from the reference lists of retrieved reports and from review articles. Unpublished reports and abstracts were not considered. Authors were not contacted for original data. Reports were included if they were randomised comparisons of peripheral opioids with

either local anaesthetics, placebo (saline), no treatment, or an opioid given by a different route, or comparisons of different doses of peripheral opioids. Reports of analgesic efficacy of intra-articular opioids are considered in chapter 16.[3]

Inclusion criteria

Each report which could possibly be described as an RCT was read independently by each of the authors and scored using a three item, 1–5 score, quality scale.[4] The scale takes into account proper randomisation, double-blinding and reporting of withdrawals and drop-outs. Consensus was then achieved. Information on doses and routes of administration of opioids and controls, types of surgery and anaesthesia, number of patients enrolled and analysed, study design, observation periods, outcome measures, and adverse effects was taken from each report.

Validity criteria

Validity criteria for included studies were number of patients per treatment group ≥ 10, any opioid except pethidine, which has shown local anaesthetic properties,[5] any peripheral site of injection except intra-articular,[3] and standardised methods of measuring sensory block and pain intensity.

Intraoperative efficacy was estimated by comparing onset and quality (loss of pinprick and touch sensation), and duration of a sensory block with opioid compared with control. Postoperative efficacy was estimated by comparing pain intensity, delay until first analgesic, and total analgesic consumption with opioid compared with control. Pain intensity measurement was analysed when reported as a VAS or verbal rating scale (VRS).

Data showing any statistically significant difference ($p < 0.05$) between opioid and control, as indicated in the original report, were extracted. The authors then met to achieve consensus (vote-counting procedure) whether such a statistically significant difference was of clinical relevance. Finally, the decision (i.e. the vote counting) reached on clinical relevance was compared with the original authors' conclusion of efficacy.

Results

In all, 45 trials were considered for analysis, of which 17 were subsequently excluded *(Table 50)*. Two further reports were not considered because no copies were available in the UK.[6,7]

TABLE 50 *Analgesic efficacy of peripheral opioids: excluded trials*

Excluded trial	Reason for exclusion
Acalovschi & Cristea, 1995[8]	Pethidine
Armstrong, *et al.,* 1993[5]	Pethidine
Davidas, *et al.,* 1992[9]	Pethidine
El Bakry, *et al.,* 1989[10]	Pethidine
Oldroyd, *et al.,* 1994[11]	Pethidine
Gobeaux & Landais, 1988[12]	Pethidine, not random
Arendt-Nielsen, *et al.,* 1990[13]	Not random
Kepplinger, *et al.,* 1995[14]	Not random
Moore, *et al.,* 1994[15]: Study 2	Number of patients per group < 10
Pere, 1993[16]	Number of patients per group < 10
Wajima, *et al.,* 1995b[17]	Number of patients per group < 10
Welte, *et al.,* 1992[18]	Number of patients per group < 10
Tenant, *et al.,* 1993[19]	Number of patients per group < 10
Ben-Ameur, *et al.,* 1993[20]	No pain outcomes
Arendt-Nielsen, *et al.,* 1991[21]: Study 2	No opioid evaluated
Bullingham, *et al.,* 1984[22]	Not analysable
Mays, *et al.,* 1987[23]	Chronic pain

Data from 26 RCTs, published in 25 reports, were analysed. In all, 952 patients, 485 of whom received an opioid, were studied *(Table 51)*.

The average size of trial was 15 patients per group (range 10–32). The median quality score was 2 (range 1–4). Three reports[31,32,34] covered four trials (8%) which included a treatment arm with an analgesic method of proven efficacy and, therefore, had an index of internal sensitivity. Eight trials (16%) used a double-dummy design.[15,32,34,36,40,42,44]

Efficacy of peripheral opioids was tested in experimental pain trials in healthy volunteers and in a wide variety of surgical settings with intravenous regional anaesthesia (Bier's block), intrapleural, intraperitoneal, incisional and dental injections, perineural blocks (femoral, ankle block, intercostal)

TABLE 51 *Analgesic efficacy of peripheral opioids: analysed RCTs*

Study	Treatments (number of patients)	Setting	Efficacy intraoperatively ('anaesthesia') (< less effective, $p < 0.05$; > more effective, $p < 0.05$; = no difference)	Efficacy postoperatively ('analgesia') (< less effective, $p < 0.05$; > more effective, $p < 0.05$; = no difference)	Adverse effects
Experimental					
Arendt-Nielsen, *et al.*, 1991[21] Study 1	1. Morphine, 4 mg, 10 ml (10) 2. Saline, 10 ml (10)	Left and right ulnar nerve block Laser stimulation	Pain and sensory thresholds and brain potentials: morphine > saline at 15 minutes only	N/A	None
Armstrong, *et al.*, 1993[5]	1. Prilocaine 0.5%, 40 ml + fentanyl, 100 µg, 2 ml (15) 2. Prilocaine 0.5%, 40 ml + saline, 2 ml (15)	Bier's block Needle and temperature stimulation	Onset, speed of recovery and quality of sensory block: fentanyl = saline	N/A	Nausea: fentanyl 7; saline 1
Arthur, *et al.*, 1992[24]	1. Lignocaine 0.5%, 100 mg, 40 ml + saline, 2 ml (10) 2. Lignocaine 0.5%, 100 mg, 40 ml + fentanyl, 100 µg, 2 ml (10) 3. Fentanyl, 100 µg, + saline, 40 ml (10)	Bier's block Needle, temperature Grip strength	Sensory and motor block (quality and onset): fentanyl < lignocaine = lignocaine + fentanyl	N/A	Nausea: lignocaine + fentanyl 2; fentanyl 1
Moniche, *et al.*, 1993[25]	1. Morphine, 2 mg, 5 ml (12) 2. Saline, 5 ml (12)	Drugs injected s.c. in the injury Burn injury (49°C on calf bilaterally)	N/A	Heat pain threshold: morphine > saline (30–330 minutes) Pressure pain threshold: morphine > saline at 30 minutes	Erythema at the site of injection: morphine 5
Bier's block					
Abdulla & Fadhil, 1992[26]	1. Lignocaine, 100 mg, 40 ml (15) 2. Lignocaine, 100 mg, + fentanyl, 50 µg (15) 3. Lignocaine, 100 mg, + pancuronium, 0.5 mg (15) 4. Lignocaine, 100 mg, + fentanyl, 50 µg, + pancuronium, 0.5 mg (15)	Upper limb surgery	VRS: (4) > (1); (4) > (3); no significant result for (2). Neuro muscular block: (3) and (4) > (1) or (2)	N/A	None
Ericyes, *et al.*, 1995[27]	1. Prilocaine 1%, 30 ml, + saline, 10 ml (10) 2. Prilocaine 1%, 30 ml, + morphine, 6 mg, 10 ml (10)	Upper limb surgery	Onset and recovery of sensory block: morphine > saline	N/A	None
Gupta, *et al.*, 1993[28]	1. Prilocaine 0.5%, 3 mg/kg, + saline, 5 ml (20) 2. Prilocaine 0.5%, 3 mg/kg, + morphine, 1 mg/5 ml (17)	Upper limb surgery	N/A	VAS PI, total analgesic consumption: morphine = saline	Mild localised urticaria: morphine 1
Pitkanen, *et al.*, 1992[29]	1. Prilocaine 0.5%, 40 ml, + saline, 4 ml, (12) 2. Prilocaine 0.5%, 40 ml, + fentanyl, 100 µg, 4 ml (13) 3. Prilocaine 0.5%, 40 ml, + fentanyl, 200 µg, 4 ml (12)	Minor surgery of upper extremity	Loss of pinprick at 15 minutes: fentanyl, 200 mg, > fentanyl, 100 mg, = saline Time to develop analgesia: fentanyl, 200 mg, = fentanyl, 100 mg, = saline	N/A	Nausea and dizziness: saline 1; fentanyl, 100 mg, 7; fentanyl, 200 mg, 6
Other peripheral sites of injection: all drugs injected postoperatively					
Aykac, *et al.*, 1995[30]	1. Morphine, 20 mg, intrapleural, 20 ml (14) 2. Morphine, 20 mg, i.v. (ml not applicable) (14)	Thoracotomy (lobectomy for lung cancer). Morphine injected after pleural closure.	N/A	VRS: intrapleural > i.v. Blood levels: i.v. > intrapleural	Confusion: intrapleural 4; i.v. 4 Urinary retention: intrapleural 3; i.v. 6 Respiratory depression: i.v. 1
Moore, *et al.*, 1994[15] Study 1	1. Morphine, 30 ng local, 0.3 ml + oral placebo (10) 2. Placebo local, 0.3 ml + morphine, 50 ng p.o. (10)	Bilateral third molar surgery; locally applied morphine.	N/A	VAS PI, analgesic consumption: morphine = saline	None
Rosenstock, *et al.*, 1996[31]	1. Morphine, 5 mg, 6 ml, incisionally (10) 2. Morphine, 5 mg, 1 ml i.v. (10) 3. Morphine, 5 mg, 6 ml s.c. (10) 4. Saline, 6 ml, incisionally (10)	Inguinal herniotomy; incisional morphine postoperatively.	N/A	VAS PI (rest, movement), analgesic consumption: morphine incision = i.v. = s.c. = saline	Not reported
Schulte-Steinberg, *et al.*, 1995[32] Study 1	1. Morphine, 1 mg, i.p. + saline i.v. (18) 2. Saline i.p. + morphine, 1 mg, i.v. (17) 3. Bupivacaine 0.25% i.p. + saline i.v. (15) NB: All drugs diluted in 20 ml.	Laparoscopic cholecystectomy; i.p. injection at the end of surgery.	N/A	VAS PI, VRS, McGill, analgesic consumption: morphine i.p. = morphine i.v. = bupivacaine i.p.	Not reported
Schulte-Steinberg, *et al.*, 1995[32] Study 2	1. Morphine, 1.5 mg, intrapleural + saline i.v. (20) 2. Saline intrapleural + morphine, 1.5 mg, i.v. (20) 3. Bupivacaine 0.25%, intrapleural + saline i.v. (20) NB: All drugs diluted in 30 ml.	Laparoscopic cholecystectomy; intrapleural injection at the end of surgery.	N/A	VAS PI, VRS, McGill, analgesic consumption: morphine intrapleural = morphine i.v. < bupivacaine intrapleural	Not reported

continued

TABLE 51 contd *Analgesic efficacy of peripheral opioids: analysed RCTs*

Study	Treatments (number of patients)	Setting	Efficacy intraoperatively ('anaesthesia') (< less effective, $p < 0.05$; > more effective, $p < 0.05$; = no difference)	Efficacy postoperatively ('analgesia') (< less effective, $p < 0.05$; > more effective, $p < 0.05$; = no difference)	Adverse effects
Perineural: drugs injected pre- or postoperatively					
Bullingham, *et al.*, 1983[33]	1. Morphine 0.02% one side + saline other side (10) 2. Morphine 0.04% one side + morphine 0.02% other side (10) NB: Volume 15–20 ml per injection	Ankle nerve block (nerves); bilateral foot minor surgery	N/A	VAS PI and PR: morphine 0.02% = morphine 0.01% = saline	Not reported
Dahl, *et al.*, 1988[34]	1. Morphine, 4 mg, epidural, 10 ml + saline femoral, 10 ml (10) 2. Saline epidural, 10 ml + morphine, 4 mg, femoral, 10 ml (10)	Femoral block and epidural catheter after knee surgery. Treatment reversed for the next 24 hours.	N/A	VAS PI: epidural > femoral Morphine consumption: epidural = femoral	Nausea, vomiting: epidural > femoral
Sternlo & Hagerdal[35]	1. Bupivacaine 0.5%, 20 ml (24) 2. Bupivacaine 0.5%, 20 ml, + morphine, 4 mg (26) NB: ml N/A	Intercostal block (4 ml per 5 ribs); biliary surgery	N/A	VAS PI, delay for analgesic, analgesic consumption: bupivacaine + morphine = bupivacaine	None
Brachial plexus					
Bourke & Furman[36]	1. Lignocaine 1.5%, 0.55 ml/kg, + morphine, 0.1 mg/kg (ml N/A) + saline i.v., 0.1 ml/kg (20) 2. Lignocaine 1.5%, 0.55 ml/kg, + morphine, 0.1 mg/kg i.v. (ml N/A) + saline, 0.1 ml/kg (20)	Hand and forearm surgery; axillary block	N/A	VAS, recovery of sensory and motor block: i.v. = axillary block Analgesic consumption: axillary block > i.v.	Mild nausea: axillary block 1; i.v. 2
Fletcher, *et al.*, 1994[37]	1. Lignocaine 1.5%, 38 ml, + fentanyl, 100 µg, 2 ml (26) 2. Lignocaine 1.5%, 38 ml, + saline, 2 ml (25)	Orthopaedic surgery; axillary block	Onset and duration of block: fentanyl = saline Success rate of each nerve block: fentanyl = saline	N/A	None
Flory, *et al.*, 1995[38]	1. Bupivacaine 0.5%, 40 ml, + morphine, 5 mg, 5 ml (20) 2. Bupivacaine 0.5%, 40 ml, +saline, 5 ml (20)	Shoulder surgery; interscalene block	N/A	VAS PI, delay until first and total dose of analgesic: morphine = saline	Nausea, vomiting: saline 5; morphine 10 Pruritus: saline 3; morphine 0 Urine retention: saline 1; morphine 1
Gobeaux & Landais[12]	1. Lignocaine 1.5%, 30 ml (12) 2. Lignocaine 1.5%, 30 ml, + fentanyl, 100 µg (12) NB: ml N/A	Upper limb surgery; axillary block	Onset and intensity of block: fentanyl > no treatment between 5 and 10 minutes	N/A	None
Gormley, *et al.*, 1996[39]	1. Lignocaine 1.5%, 7 mg/kg, + alfentail, 10 µg /kg, 10 ml (28) 2. Lignocaine 1.5%, 7 mg/kg, + saline 10 ml (32)	Upper limb surgery; axillary block	Duration of sensory and motor block: alfentanil > saline (10–40 minutes)	VAS: alfentanil > saline (h 3) Delay for first analgesic, block recovery: alfentanil = saline	None
Kardash, *et al.*, 1995[40]	1. Mepivacaine 1.5%, 30 ml, + fentanyl, 75 µg, 1.5 ml, + saline i.m., 1.5.ml (10) 2. Mepivacaine 1.5%, 30 ml,+ fentanyl, 75 µg i.m., 1.5 ml, + saline, 1.5.ml (10)	Upper limb surgery; supraclavicular block	Onset, duration of sensory and motor block: axillary block = i.m.	VAS PI axillary block > i.m. (0–1 hours)	None
Morros Vinoles, *et al.*, 1991[41]	1. Mepivacaine 1%, 40 ml 2. Mepivacaine 1%, 40 ml, + fentanyl, 100 µg (ml N/A) 3. Mepivacaine 1%, 40 ml, + fentanyl, 100 µg s.c.	Upper limb surgery; axillary block	Onset and quality of block: (1) = (2) = (3)	Delay until first analgesia: (1) = (2) = (3)	None
Racz, *et al.*, 1991[42]	1. Bupivacaine 0.5% + lignocaine 1%, 40 ml, + morphine, 5 mg, 1 ml (19) 2. Bupivacaine 0.5% + lignocaine 1%, 40 ml, + morphine, 5 mg i.m., 1 ml (21)	Arm and forearm minor surgery; axillary block	Onset and quality of sensory and motor block: axillary block = i.m.	VAS PI, delay until first analgesic: axillary block = i.m.	None
Viel, *et al.*, 1989[43]	1. Bupivacaine 0.5%, 40 ml, + morphine, 50 µg/kg (20) 2. Bupivacaine 0.5%, 40 ml, + buprenorphine, 3 µg/kg (20)	Upper limb surgery; supraclavicular block	Sensory and motor block: buprenorphine = morphine	Quality, duration of analgesia: buprenorphine > morphine	Pruritus: morphine 1 Nausea, vomiting: buprenorphine 1
Wajima, *et al.*, 1995[44]	1. Butorphanol, 83 µg/h + saline i.v. (12) 2. Butorphanol, 83 µg/h i.v. + saline (10) NB: all perfusion 50 ml/72 hours)	Upper extremity surgery; axillary block (postoperatively continuous infusion)	N/A	VAS PI at 9, 12, 18, 20 hours: axillary block > i.v. Supplementary analgesia: i.v. = axillary block	Nausea: i.v. 6; axillary 4 Vomiting: i.v. 2; axillary 2 Drowsiness: i.v. 1; axillary 2

and brachial plexus sheath injections (axillary, supraclavicular and interscalene approaches). The opioids used were morphine (16 trials), fentanyl (8), alfentanil (1), buprenorphine (1) and butorphanol (1). Intraoperative efficacy assessments were performed in ten clinical trials. Postoperative efficacy was evaluated in 17 clinical trials.

Experimental pain trials (4 trials)

In one trial, morphine was applied perineurally;[21] sensory and pain thresholds were significantly increased compared with saline but for not longer than 15 minutes. One trial used morphine subcutaneously at the site of injury and reported higher heat and pain thresholds compared with saline.[25] Two other trials failed to demonstrate any benefit from adding fentanyl to a local anaesthetic in a Bier's block.[24,45]

The experimental nature of these reports makes it difficult to judge clinical relevance. Therefore they were not taken into account in estimating overall efficacy of peripheral opioids.

Bier's block (4 trials)

Fentanyl was used in two trials.[26,29] Abdulla and Fadhil could not demonstrate any significant difference between the combination of fentanyl plus local anaesthetic and local anaesthetic alone,[26] but nevertheless concluded that the method was of clinical relevance. We disagreed with these authors because they did not comment on the comparison of interest to us (i.e. opioid versus no treatment) but rather based their conclusion on the comparison between opioid plus curare versus no treatment or curare alone.

Pitkånen and colleagues reported a significantly improved quality of the sensory block after 15 minutes with fentanyl, 200 µg, compared with either saline or fentanyl, 100 µg.[29] No measurements were taken after 15 minutes. Nausea and dizziness were more frequent with fentanyl. These authors concluded that their finding was not clinically relevant.

Morphine was used in two trials.[27,28] In one there was no significant difference between morphine and saline, and the authors concluded that morphine was of no value in Bier's block.[28] In the other trial, both onset of and recovery from anaesthesia and analgesia were significantly better with morphine compared with local anaesthetic alone.[27] These authors concluded that the differences of 1 minute and 2 minutes, respectively, were clinically relevant. We disagreed, because the difference between the two groups was of very short duration only and therefore of no practical importance.

Other peripheral sites (5 trials)

All five trials used morphine. Four of them could not demonstrate any difference in the postoperative period between morphine and control when applied into a tooth socket,[15] into a surgical wound,[31] or by intraperitoneal or intrapleural block (study I and II).[32] The fifth trial reported a statistically significant improvement with morphine, 20 mg, given intrapleurally compared with the same drug and dose given intravenously;[30] verbal pain rating scores were lower for 20 hours in the intrapleural group. Morphine plasma levels were lower in the intrapleural group. This was considered to be of clinical relevance by these authors. Analgesic consumption was not reported. We considered the outcome to be of little clinical relevance because of the unconventional (high) dose of morphine used.

Perineural (3 trials)

None of these trials reported any significant difference between the opioid and control.[33–35]

Brachial plexus (ten trials)

Opioids were given by interscalene (one trial), supraclavicular (2) or axillary (7) approaches to the brachial plexus sheath.

In three trials, morphine was combined with a local anaesthetic and applied either by axillary[36,42] or interscalene route.[38] Comparators were systemic morphine or axillary saline. No intraoperative or postoperative improvement could be demonstrated with peripheral morphine in two of the three trials.[38,42] The third trial (axillary route) reported similar pain scores in the groups but a significantly lower postoperative analgesic consumption (number of tablets of oxycodone, 5 mg, plus acetaminophen, 500 mg) with the opioid; the authors concluded that this difference was clinically important.[36] The median number of tablets was two with axillary morphine and four with systemic morphine. We did not consider this difference to be of clinical importance in this acute setting.

In four trials, fentanyl was combined with a local anaesthetic and compared with a local anaesthetic alone or with another route of injection.[37,40,41,46] In two a significant improvement with fentanyl was reported.[40,46] Gobeaux and colleagues concluded that a faster speed of onset of the sensory block with the opioid was clinically relevant.[46] However, this difference was only 5 minutes. Kardash and colleagues reported a lower VAS for pain intensity for the first postoperative hour with fentanyl but did not consider this to be clinically important.[40]

Alfentanil added to a local anaesthetic led to a significant improvement compared with the local anaesthetic plus placebo;[39] duration of sensory and motor block after surgery was 40 minutes longer with the opioid. This was considered to be clinically relevant by these authors, although there was no difference between the two groups in the delay until the first analgesic rescue medication.

Butorphanol perfusion into the plexus sheath led to significantly lower VAS scores for pain intensity up to hour 24 postoperatively compared with the same butorphanol perfusion given intravenously.[44] There was no difference in postoperative analgesic requirements. The authors concluded that this difference was clinically relevant. However, average VAS scores were very low, irrespective of the route of administration (i.e. axillary route 6–7% of the maximum on a VAS for pain intensity; intravenous route 17–33%).

Buprenorphine, 3 µg/kg, was compared with morphine, 50 µg/kg, in one trial; both opioids were added to the same local anaesthetic before supraclavicular injection.[43] A placebo group was lacking in this trial. Duration and quality of postoperative analgesia were significantly better with buprenorphine; 'good' pain relief, as judged by the patients, lasted for 35 hours with buprenorphine and 18 hours with morphine. Authors concluded that buprenorphine is efficacious and long-acting as an analgesic when injected into the brachial plexus sheath. However, they did not take into account equi-analgesic dosing.

There was a relationship between quality scores of the reports and original authors' conclusions on efficacy of peripheral opioids *(Figure 45)*. Authors of ten trials (two experimental and eight clinical) reported positive estimates of efficacy. Quality scores for these trials were 2 or below.[21,25–27,30,36,39,43,44,46] In the 16 remaining trials (two experimental and 14 clinical) conclusions were negative. Seven, including one with two studies,[32] had a quality score of 3 or 4.[28,37,38,42,45]

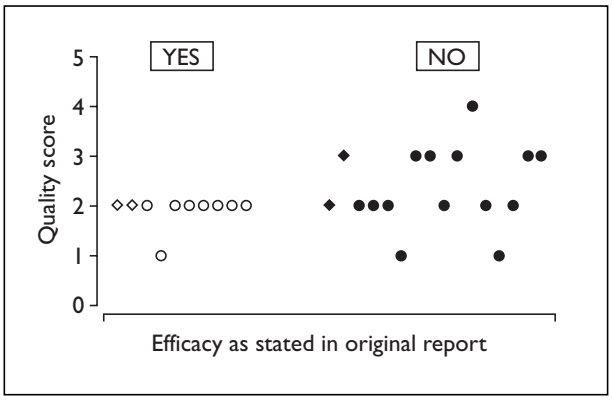


FIGURE 45 *Efficacy of peripheral opioids (relationship between quality of trials and overall examination of efficacy as stated in original reports). ○, ●, clinical trials; ◇, ◆, experimental pain trials.*

Adverse effects

No adverse effects attributable to the route of administration were reported.

Comment

The aim of this systematic review was to test the evidence for an analgesic action of peripheral opioids and the clinical relevance of such action. In all, 26 RCTs with data from more than 950 patients were analysed. These trials described a variety of surgical procedures and experimental designs. Five different opioids with several different doses were administered with ten different regional anaesthetic techniques. Trials were not consistent in either analysing or reporting quality of surgical blocks, postoperative analgesia, or observation periods. Estimation of efficacy based on data of such methodological and clinical heterogeneity was, therefore, difficult. Quantitative analysis was impossible. Unfortunately the different procedures or blocks operate, or may operate, in different ways, so that a negative result from one procedure does not preclude a positive result with another.

It was decided to judge the studies in two ways: those which had any result which was statistically in favour of a peripheral action of opioid, and those where the result was of a sufficient magnitude or importance to be clinically relevant. Because most of the studies had a number of different measurements at different times, the possibility that some statistical differences could occur by chance was high. Use of the conventional level of statistical significance in clinical and scientific studies of the 95th percentile implies that, if 20 different measurements are made, one will show significance just by chance. So, in 26 different studies with a large number of outcome measurements, some statistically significant differences with opioid would be expected. Judging clinical relevance may be easier or more difficult. Most practising clinicians would claim they could tell when a result was going to benefit their patients. Codifying what that entails is not easy. Reaching a consensus on clinical efficacy of peripheral opioids was influenced by whether all the measures in a study showed statistically significant differences, or whether the magnitude of any difference shown in a trial was sufficiently large to make change in

practice a reasonable consideration. The authors are conscious that others might reach different conclusions.

Of 26 trials, 14 were unequivocally negative. The remaining 12 trials[21,25–27,29,30,36,39,40,43,44,46] reported at least one statistically significant result in favour of the peripheral opioid. Of these 12 positive trials, two were in experimental pain;[21,25] their results may not be directly applicable to clinical practice. Of the remaining ten positive trials, authors of two[29,40] did not regard their findings as being clinically relevant. This means that results from only eight out of the 26 trials (31%) were judged by their original authors as clinically relevant.

We could not, for different reasons, support the conclusions of any of these positive reports. An isolated significant outcome in favour of the opioid, such as a longer duration of a sensory block which was not correlated with a delay until the patient needed a first analgesic, was not judged clinically relevant by us.[39] Differences of doubtful clinical importance were reported, such as the shortening of the onset of a surgical block by a few minutes[27,46] or a minimal difference in the average analgesic consumption.[36] A significant difference between two opioids was shown but without a placebo control,[43] so that clinical relevance of this greater efficacy of buprenorphine relative to morphine remains questionable. Furthermore, in this trial the two opioids were compared in non-equi-analgesic doses, and a systemic analgesic effect of buprenorphine with its long duration of action cannot be ruled out. Other drawbacks in studies with positive findings were the very low pain intensity scores irrespective of the treatment,[44] the unconventional dose of opioids used,[30] or the comparison of treatment arms which were of no interest to this review.[26] Such trials cannot be regarded as valid assays for evaluation of analgesic efficacy.

Do these trials represent evidence of a lack of efficacy of peripheral opioids, or rather a lack of evidence of their efficacy? In the systematic review of the relevant published literature on the analgesic efficacy of intra-articular morphine, only a minority of the analysed data could be regarded as valid; yet this limited amount of data provided some evidence for its analgesic efficacy.[3] Validity in those trials was assumed when baseline pain was sufficiently high to allow measurement of pain relief, when an index of internal sensitivity was given, and when blinding was adequate.[3] Most of the trials in the present review did not meet these criteria.

As well as the issue of validity, there is the issue of methodological quality. There are other examples where trials with low scores (two or below on a scale of 1–5) on the validated quality scale[4] used in this review, have overestimated the effectiveness of treatment.[47] In the present review, none of the ten trials which claimed efficacy of peripheral opioids had a score above two *(Figure 45)*. Seven of the 14 unequivocally negative reports had scores of three or four. This means that the trials of highest methodological quality in this data set could not show any difference between peripheral opioid and control.

This subgroup analysis by trial quality emphasises that, in these clinical models, peripheral opioids have no efficacy. The question is then why good quality trials showed some efficacy of morphine in the knee joint but no efficacy of different opioids in peripheral sites outside the knee joint. This may be because the knee joint model better reflects the inflammatory process which is thought to be of importance in sensitising peripheral opioid receptors.[48] It may also be related to inadequately low doses of opioids used in these trials. Doses of morphine between 0.5 mg and 5 mg tested in the confined space of the knee joint produced very high local concentrations.[3] Similar doses of morphine injected into the peritoneal or pleural cavity, or into an isolated limb would produce much lower local concentrations than in the knee joint.

This qualitative analysis of pain trials highlights the importance of critical appraisal of the literature and some of the difficulty encountered in doing it. Authors of original reports tended to overinterpret their findings and to confuse statistical significance with clinical relevance. Inattentive or uncritical readers may be misled into a false perception of treatment efficacy. Some 30 years ago, Schwartz and Lellouch distinguished between explanatory studies, designed to prove a hypothesis, and pragmatic studies, designed to tell us whether instigating a change was of benefit.[49] The distinction is still important and the clinical use of peripheral opioids requires much more evidence than exists at present.

References

1. Picard PR, Tramèr MR, McQuay HJ, Moore RA. Analgesic efficacy of peripheral opioids (all except intra-articular): a qualitative systematic review of randomised controlled trials. *Pain* 1997;**72**: 309–18.

2. Stein C. The control of pain in peripheral tissue by opioids. *N Engl J Med* 1995;**332**:1685–90.

3. Kalso E, Tramèr M, Carroll D, McQuay H, Moore RA. Pain relief from intra-articular morphine after knee surgery: a qualitative systematic review. *Pain* 1997;**71**:642–51.

4. Jadad AR, Moore RA, Carroll D, Jenkinson C, Reynolds DJM, Gavaghan DJ, *et al.* Assessing the quality of reports of randomized clinical trials: is blinding necessary? *Control Clin Trials* 1996;**17**:1–12.

5. Armstrong PJ, Morton CPJ, Nimmo AF. Pethidine has a local anaesthetic action on peripheral nerves in vivo. *Anaesthesia* 1993;**48**:382–6.

6. Mocavero G. Analgesia selettiva con morfina perinervosa. *Incontri Anest Rianim Sci Affini* 1981;**16**:1–3.

7. Vinoles CM, Cuenca P, Monsegur JC, Valloba FC. Adicion de fentanilo a la mepivacaina en el bloqueo axilar del plexo braquial. *Rev Esp Anestiol Reanim* 1991;**38**:87–9.

8. Acalovschi I, Cristea T. Intravenous regional anesthesia with meperidine. *Anesth Analg* 1995;**81**:539–43.

9. Davidas JL, Blond JL, Rochette A, Manchon M, Degoute CS, Bansillon V. Action analgésique locale par effet directe sur les troncs nerveux de la péthidine. *Thérapie* 1992;**47**:485–7.

10. El Bakry MS, El-Shafei SB, Seyam EM, El-Kobbia NM, Ebrahim UH. Use of pethidine as an intravenous regional anesthetic. *Middle East J Anesth* 1989;**10**:189–94.

11. Oldroyd GJ, Tham EJ, Power I. An investigation of the local anaesthetic effects of pethidine in volunteers. *Anaesthesia* 1994;**49**:503–6.

12. Gobeaux D, Landais A. Utilisation de deux morphiniques dans les blocs du plexus brachial. *Cah Anesthesiol* 1988;**36**:437–40.

13. Arendt-Nielsen L, Oberg B, Bjerring P. Laser-induced pain for quantitative comparison of intravenous regional anesthesia using saline, morphine, lidocaine, or prilocaine. *Reg Anesth* 1990;**15**:186–93.

14. Kepplinger B, Schmid H, Rettensteiner G, Derfler C, Papst H, Erhart H, *et al.* Wirkungsvergleich zwischen intramuskulär und periradikuär verabreichtem Tramadol. *Neuropsychiatrie* 1995;**9**:196–200.

15. Moore UJ, Seymour A, Gilroy J, Rawlins MD. The efficacy of locally applied morphine in post-operative pain after bilateral third molar surgery. *Br J Clin Pharmacol* 1994;**37**:227–30.

16. Pere P. The effect of continuous interscalene brachial plexus block with 0.125% bupivacaine plus fentanyl on diaphragmatic motility and ventilatory function. *Reg Anesth* 1993;**18**:93–7.

17. Wajima Z, Shitara T, Nakajima Y, Kim C, Kobayashi N, Kadotani H, *et al.* Comparison of continuous brachial plexus infusion of butorphanol, mepivacaine and mepivacaine-butorphanol mixtures for postoperative analgesia. *Br J Anaesth* 1995;**75**:548–51.

18. Welte M, Haimerl E, Groh J, Briegel J, Sunder-Plassmann L, Herz A, *et al.* Effect of intrapleural morphine on postoperative pain and pulmonary function after thoracotomy. *Br J Anaesth* 1992;**69**:637–9.

19. Tenant F, Moll D, de Paulo V. Topical morphine for peripheral pain. *Lancet* 1993;**342**:1047–8.

20. Ben-Ameur M, Ecoffey C, Kuhlman G, Mazoit X, Gobeaux D. Réponse ventilatoire au dioxyde de carbone après bloc axillaire avec fentanyl et lidocaïne. *Ann Fr Anesth* 1993;**12**:22–6.

21. Arendt-Nielsen L, Bjerring P, Berg Dahl J. A quantitative double-blind evaluation of the antinociceptive effects of perineurally administered morphine compared with lidocaine. *Acta Anesth Scand* 1991;**35**:24–9.

22. Bullingham RE, McQuay HJ, Moore RA. Studies on the peripheral action of opioids in postoperative pain in man. *Acta Anaesthesiol Belg* 1984;**35** suppl: 285–90.

23. Mays KT, Lipman JJ, Schnapp M. Local analgesia without anesthesia using peripheral, perineural morphine injections. *Anesth Analg* 1987;**66**:417–20.

24. Arthur JM, Heavner JE, Tanmian MB, Rosenberg PH. Fentanyl and lidocaine versus lidocaine for Bier block. *Reg Anesth* 1992;**17**:223–7.

25. Moniche S, Dahl JB, Kehlet H. Peripheral antinociceptive effects of morphine after burn injury. *Acta Anaesth Scand* 1993;**37**:710–12.

26. Abdulla WY, Fadhil NM. A new approach to intravenous regional anesthesia. *Anesth Analg* 1992;**75**:597–601.

27. Erciyes N, Akturk G, Solak M, Dohman D. Morphine/prilocaine combination for intravenous regional anesthesia. *Acta Anaesth Scand* 1995;**39**:845–6.

28. Gupta A, Bjornsson A, Sjoberg F, Bengtsson M. Lack of peripheral analgesic effect of low dose morphine during intravenous regional anesthesia. *Reg Anesth* 1993;**18**:250–3.

29. Pitkånen MT, Rosenberg PH, Pere PJ, Tuominen MK, Seppala TA. Fentanyl–prilocaine mixture for intravenous regional anaesthesia in patients undergoing surgery. *Anaesthesia* 1992;**47**:395–8.

30. Aykac B, Erolcay H, Dikmen Y, Oz H, Yillar O. Comparison of intrapleural versus intravenous morphine for post thoracotomy pain management. *Cardiothor Vasc Anesth* 1995;**9**:538–40.

31. Rosenstock C, Andersen G, Antonsen K, Rasmussen H, Lund C. Analgesic effect of incisional morphine following inguinal herniotomy under spinal anesthesia. *Reg Anesth* 1996;**21**:93–8.

32. Schulte-Steinberg H, Weninger E, Jokisch D, Hofstetter B, Misera A, Lange V, *et al.* Intraperitoneal versus interpleural morphine or bupivacaine for pain after laparoscopic cholecystectomy. *Anesthesiology* 1995;**82**:634–40.

33. Bullingham R, O'Sullivan G, McQuay H, Poppleton P, Rolfe M, Evans P, *et al.* Perineural injection of morphine fails to relieve postoperative pain in humans. *Anesth Analg* 1983;**62**:164–7.

34. Dahl JB, Daugaard JJ, Kristoffersen E, Johannsen HV, Dahl JA. Perineuronal morphine: a comparison with epidural morphine. *Anaesthesia* 1988;**43**:463–5.

35. Sternlo IE, Hagerdal M. Perineuronal morphine in intercostal block. *Anaesthesia* 1992;**47**:613–15.

36. Bourke DL, Furman WR. Improved postoperative analgesia with morphine added to axillary block solution. *J Clin Anesth* 1993;**5**:114–17.

37. Fletcher D, Kuhlman G, Samii K. Addition of fentanyl to 1.5% lidocaine does not increase the success of axillary plexus block. *Reg Anesth* 1994;**19**:183–8.

38. Flory N, Van Gessel E, Donald F, Hoffmeyer P, Gamulin Z. Does the addition of morphine to brachial plexus block improve analgesia after shoulder surgery? *Br J Anaesth* 1995;**75**:23–6.

39. Gormley WP, Murray JM, Fee JPH, Bower S. Effect of the addition of alfentanil to lignocaine during axillary brachial plexus anaesthesia. *Br J Anaesth* 1996;**76**:802–5.

40. Kardash K, Schools A, Concepcion M. Effects of brachial plexus fentanyl on supraclavicular block: a randomized, double-blind study. *Reg Anesth* 1995;**20**:311–15.

41. Morros Vinoles C, Perez Cuenca MD, Castillo Monsegur J, Cedo Valloba F. Adicion de fentanilo a la mepivacaina en el bloqueo axilar del plexo braquial. *Rev Esp Anestesiol Reanim* 1991;**38**:87–9.

42. Racz H, Gunning K, Della Santa D, Forster A. Evaluation of the effect of perineuronal morphine on the quality of postoperative analgesia after axillary plexus block: a randomised double-blind study. *Anesth Analg* 1991;**72**:769–72.

43. Viel EJ, Eledjam JJ, de La Coussaye JE, d'Athis F. Brachial plexus block with opioids for postoperative pain relief: comparison between buprenorphine and morphine. *Reg Anesth* 1989;**14**:274–8.

44. Wajima Z, Nakajima Y, Kim C, Kobayashi N, Kadotami H, Adachi H, *et al.* I.v. compared with brachial plexus infusion of butorphanol for postoperative analgesia. *Br J Anaesth* 1995;**74**:392–5.

45. Armstrong PJ, Power I, Wildsmith JAW. Addition of fentanyl to prilocaine for intravenous regional anesthesia. *Anaesthesia* 1991;**46**:278–80.

46. Gobeaux D, Landais A, Bexon G, Cazaban J, Levron JC. Adjonction de fentanyl à la lidocaéne adrénalinée pour le blocage du plexus brachial. *Cah Anesthesiol* 1987;**35**:195–9.

47. Khan KS, Daya S, Jadad AR. The importance of quality of primary studies in producing unbiased systematic reviews. *Arch Intern Med* 1996;**156**:661–6.

48. Nagasaka H, Awad H, Yaksh T. Peripheral and spinal actions of opioids in the blockade of the autonomic response evoked by compression of the inflamed knee joint. *Anesthesiology* 1996;**85**:808–16.

49. Schwartz D, Lellouch J. Explanatory and pragmatic attitudes in therapeutic trials. *J Chronic Dis* 1967;**20**:637–48.

Chapter 18

Pre-emptive analgesia: a systematic review of clinical studies to 1994

Summary

Basic science evidence suggests that an analgesic intervention made before surgery will produce a better outcome than the same intervention made after surgery. The evidence from RCTs which tested this hypothesis in patients is reviewed in this chapter.

Four studies with paracetamol or NSAIDs did not show any pre-emptive effect. Of seven studies with local anaesthetic, six did not show a pre-emptive effect. In four studies with opioids, there was weak evidence of a pre-emptive effect in three.

There are few perfect RCTs and, unfortunately, this applies in the pre-emptive analgesia field. Many of the studies which did not show a pre-emptive effect lacked power. The opioid studies which did show a pre-emptive effect had other technical weaknesses.

One way to combat lack of power would be to combine data (meta-analysis). This is very difficult in this field because of the outcome measures which investigators are using.

Introduction

Pre-emptive analgesia is analgesia given before the painful stimulus begins. The reason for giving analgesia before the painful stimulus is to prevent or reduce subsequent pain. The concept that pre-emptive analgesia might provide better pain control came from basic science studies. Initial observations were that noxious stimuli induced changes in neural function,[1] such as hyperexcitability, in the spinal cord. Later studies suggested that analgesia given before the nociceptive stimulus began was more effective than the same dose given after the stimulus.

The editorial by Wall[2] focused clinicians' attention on pre-emptive analgesia and linked fundamental work to clinical studies. He related the findings in fundamental studies, the ways in which the central nervous system changed following nociceptive stimuli and the methods which could pre-empt these changes, to clinical management of postoperative pain. Since that editorial was published, the issues have become much more focused.

The central question is whether an intervention made before pain starts has greater analgesic effect than the same intervention (same dose, same route) made after the pain. The aim is to define the questions that need to be asked and, by reviewing the clinical evidence systematically, to see whether or not definitive answers exist. This is a very active area of clinical research, so that any conclusions may be overtaken by new evidence.

The concept is a simple one. The effect of the pre-emptive analgesia is to prevent or reduce the development of any 'memory' of the pain stimulus in the nervous system. Preventing or reducing the pain memory should lower any subsequent analgesic needs.[3] The scientific interest in this phenomenon is in the underlying mechanism. The clinical interest is in the potential for improving postoperative pain management.

The concept, and the explanation, are very attractive. Management of postoperative pain has rightly been criticised many times over the last 30 years. Despite the advent of increasingly high-technology approaches, it is doubtful that most patients are any better served. If pre-emptive analgesia worked then these patients' pain might be reduced. Unfortunately there are difficulties with the details, and difficulties in interpreting conflicting evidence from clinical studies.

We have been slow in distinguishing that pre-emptive treatment with one kind of analgesic intervention, for instance opioids, may not give the same answer as pre-emptive treatment with another, such as NSAIDs. We have also been slow to distinguish between two very different outcomes, the outcome of a pre-emptive treatment on nociceptive pain and the outcome of a pre-emptive treatment on neuropathic pain.

Problems with the fundamental evidence

Timing is one critical problem. If pre-emptive treatment reduces the memory of the subsequent noxious stimulus, how long does this effect last? Evidence of any pre-emptive effect is of great academic interest but a very short-lived effect, less than 2 hours, for example, might be of little clinical relevance, particularly if the pre-emptive treatment carried any risk of increased morbidity. Conversely, pre-emptive treatments which lasted for 10 hours with minimal increase in morbidity would be of immense clinical importance. Extrapolating from brief effects demonstrated in various animal models to clinical pain is not easy.

A second problem is whether any pre-emptive effect is an effect on acute postoperative pain (nociceptive pain), or on the development of long-term sequelae such as phantom limb pain (neuropathic pain), or on both. Different pre-emptive interventions might be required to tackle these two different problems. Positive or negative evidence of an effect of a particular intervention on nociceptive pain might not apply to neuropathic pain, and vice versa.

One animal model in which pre-emptive analgesic effects have been shown is the formalin test. Subcutaneous injection of formalin into the paw gives rise to two 'peaks' of nociceptive input. Interventions may be made at various times relative to the injection of formalin, and the relative efficacy of the same intervention made before the formalin injection may be compared with the same injection made after the formalin injection. With opioids, intrathecal injection of the enkephalin DAMGO **before** the formalin produced 70% greater inhibition of the C-fibre response than the same dose injected intrathecally **after** the formalin.[4,5] With peripheral infiltrations of local anaesthetic, one before and one after the formalin injection, the behavioural response to the formalin was abolished.[6] Infiltration with local anaesthetic 25 minutes **after** the formalin made the hindpaw anaesthetic but did not abolish the behavioural response. Intrathecal injection of local anaesthetic **before** the formalin abolished the behavioural response; the same intrathecal dose 5 minutes **after** the formalin had no effect.[6] NSAIDs injected systemically or intrathecally **before** the formalin injection produce a reduction in the behavioural response;[6,7] it is not clear whether the same dose of NSAID given **after** the formalin is less effective.

The 'end' of the second peak of the formalin model occurs within an hour. This is very brief when compared with clinical pain. In another animal model, however, the development of autotomy after peripheral nerve section, longer-term 'pre-emptive' effects have been reported with the use of local anaesthetic. The speed with which autotomy developed in response to nerve section, and the severity of the autotomy, was altered by applying local anaesthetic to the nerve fibre before the operation.[8] Pre-emptive use of local anaesthetic delayed the onset of autotomy (42 versus 23 days) and reduced its severity (15 versus 41%). Similarly, 50 µg of intrathecal morphine reduced autotomy following unilateral sciatic nerve section.[9] These studies are perhaps more analogous to chronic rather than acute pain. The effect of such nerve injury is believed to be analogous to neuropathic as opposed to nociceptive pain.

From basic science then comes the idea that the same dose of an analgesic given by the same route may be more effective if given **before** surgery rather than **after**. Neither of these models operates on a time-scale which is a totally convincing analogy of the clinical operative and postoperative states. The formalin model is perhaps too brief and the autotomy model too long. The formalin model involves inflammatory change, the autotomy model nerve damage. Clinical procedures may involve both inflammatory response and nerve damage. Could pre-emptive analgesia alter outcome in all pain contexts, or is it limited, operating in, for example, somatic but not visceral pain, and are the underlying mechanisms the same? Clinical demonstration of pre-emptive analgesia might fail if the wrong setting was chosen. The secondary issue is which is the pertinent outcome? For clinical postoperative pain the outcome is measured over hours extending to days. The prevention of chronic pain development requires outcome measurement over weeks, months and, perhaps, years.

The clinical evidence

The aim of this study is to provide a systematic review of the evidence that an intervention given before the pain starts has greater effect than the same intervention (same dose, same route) given after the pain. The review is performed separately for each of three classes of intervention, NSAIDs, local anaesthetics and opioids.

The inclusion criterion for the review was RCTs which addressed the question of pre-emptive treatment versus the same treatment given after the pain had begun *(Figure 46)*. Randomised studies reduce the chance of selection bias; studies which are not randomised have no such protection. Ideally, the studies should be double-blind, and

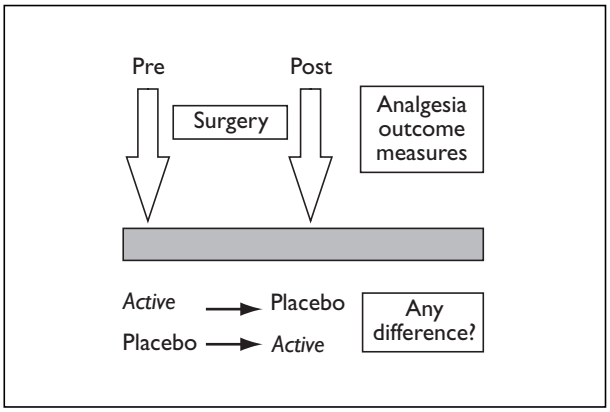


FIGURE 46 *Design of study required to show pre-emptive effect*

also double-dummy if different routes are to be compared in the treatment and control groups. Studies were excluded from this review if they were not RCTs and if they were RCTs which did not compare pre-emptive with post treatment *(Figure 46)*.

Studies were identified by a MEDLINE search and by hand searching. The MEDLINE search (Silver Platter MEDLINE v. 3.0 and 3.1) covered 1966–May 1993. The strategy was designed to identify the maximum number of randomised and/or double-blind reports by using a combination of text words, 'wild cards' and MeSH terms as described previously.[10] Medical journals were searched by hand. They were selected from a list of the 50 journals with the highest number of reports in MEDLINE, and nine specialist journals which were not included in that list or which were not indexed. The search process included volumes published between 1950 and 1994. The studies included (and excluded) are shown in *Table 52*.[11–40]

Excluded studies

Comparisons of pre-emptive treatment with no treatment (whether or not randomised)

Several of the excluded papers *(Table 52)* are often quoted as showing evidence of a pre-emptive effect. They were, however, designed to show that an analgesic intervention made **before** surgery was more effective than no intervention at all, and did not ask whether an analgesic intervention made before surgery is more, less or as effective than the same intervention made **after** surgery. Nonetheless, reviewers have reached the conclusion that these studies produce evidence of a pre-emptive effect. While a positive result in such studies suggests a worthwhile clinical benefit, it is not evidence for or against a pre-emptive effect. Such evidence requires the control of the same intervention made after surgery *(Figure 46)*.

Comparisons of pre-emptive treatment with pre-emptive plus post-treatment

These studies (all were NSAID studies) were designed to compare an NSAID given before surgery with the same NSAID given both before and after surgery. It is not possible from these studies to answer the question of whether an analgesic intervention made before surgery is more effective, less effective or as effective as the same intervention made after surgery.

Studies included

NSAID and paracetamol

Three RCTs with NSAIDs and one with paracetamol met the inclusion criterion. All were in oral surgery patients. Flath and colleagues[11] studied four groups of 30 patients each having endodontic surgery *(Table 53)*. One of these four groups had pre-operative flurbiprofen and postoperative placebo. A second group had preoperative placebo and postoperative flurbiprofen. The preoperative dose was given 30 minutes before surgery and the post-operative dose 3 hours after. On the outcome measures of categorical scale of pain intensity and VAS pain intensity, there was no evidence of a pre-emptive effect. The study had adequate sensitivity to detect a difference because one of the groups had preoperative and postoperative placebo, and the pain scores in that group were significantly higher than those in the groups who had flurbiprofen.

Sisk and colleagues[14] compared diflunisal, 1 g, with placebo in 20 patients having third molar extractions *(Table 53)*. The design was crossover, as shown in *Figure 46*. Over 8 hours there was no significant difference between preoperative and postoperative dosing, using categorical and VAS pain intensity scores. Sisk and Grover[17] used a similar design to investigate naproxen, 550 mg, in third molar extraction *(Table 53)*. Again there was no significant difference between pre- and postoperative dosing, using categorical and VAS pain intensity scores.

Gustafsson and colleagues[20] also used a two-occasion crossover design in third molar extraction, comparing paracetamol, 1 g, with placebo *(Table 52)*. Using a VAS pain intensity scale and time to first analgesic as outcome measures there was no significant difference between preoperative and postoperative dosing.

These four studies provide a consistent answer to the question. No measurable difference was found between the same dose given preoperatively and postoperatively. All four studies necessarily used local anaesthetic; none used opioids. The balance of the evidence is therefore that, at normal

therapeutic oral doses of NSAID, no pre-emptive effect was demonstrable.

Local anaesthetic

Studies of pre-emptive effect with local anaesthetics may be divided into trials of epidural (spinal), nerve block and infiltration *(Table 54)*.

Epidural

Dahl and colleagues[12] used a parallel group design on 32 colonic surgery patients *(Table 54)*. Epidural bolus and infusion of a local anaesthetic and opioid combination were given 40 minutes before surgery for the preoperative group, and after surgery for the postoperative group (some 2 hours after the preoperative group). There was no dummy injection. On categorical and VAS pain intensity scales there was no evidence of a pre-emptive effect.

Pryle and colleagues[15] used a similar study design in 36 abdominal hysterectomy patients *(Table 54)*. Local anaesthetic with adrenaline was given as a lumbar epidural bolus either 40 minutes before incision or after surgery (75 minutes after the preoperative group). On the outcome measures of categorical and VAS pain intensity scales, time to first use of intravenous morphine via PCA and amount of intravenous morphine via PCA, there was no demonstrable pre-emptive effect.

Rice and colleagues[21] compared caudal blocks pre- and postsurgery *(Table 54)* in 40 children having outpatient surgery (mean operation time 30 minutes). An objective pain score did not show any pre-emptive effect. Gunter and colleagues[23] used a similar design in 24 boys having hypospadias repair *(Table 54)*. The caudal block before surgery did reduce operating time and blood loss significantly compared with the same block after surgery, but there was no significant difference in the pain outcomes of time to first analgesic or on overall analgesic consumption.

Nerve block

Dierking and colleagues[18] compared inguinal field block pre- and postoperatively in 32 patients having herniorrhaphy *(Table 54)*. Using categorical and VAS pain intensity scales there was no evidence of a pre-emptive effect.

Infiltration

Ejlersen and colleagues[24] investigated pre- and postoperative wound infiltration in 37 herniorrhaphy patients *(Table 54)*. Using time to remedication as the outcome measure, the patients who had the infiltration 5 minutes before incision had significantly longer time until remedication, clear evidence of a pre-emptive effect.

Turner and Chalkiadis[25] compared infiltration after induction (29 patients) with infiltration after surgery (32 patients) and with no infiltration (29 patients) in appendicectomy *(Table 54)*. The outcome measures were VAS pain intensity scores and PCA consumption. They found no significant difference between the groups.

The study which did show a pre-emptive effect is the least subject to criticism,[24] but it is still balanced by six negative studies, one of which is an infiltration study of similar design. The paediatric studies[21,23] both had the problem of pain scoring in children but neither showed any pre-emptive effect. Importantly neither study involved the use of opioid; both studies therefore sought but did not find a pre-emptive effect of local anaesthetic alone. It is difficult to understand why one infiltration study should have produced a positive result[24] when the other did not.[25]

In all the negative studies the local anaesthetic intervention worked well, whether given before or after surgery. The power of these studies (is the lack of difference a true result?) thus becomes a major issue. Also, it is not known whether these studies were sufficiently sensitive to measure an effect if there was one. The one study which did use a no-treatment control[25] did not measure any significant difference between infiltration (pre or post) and no infiltration. This presents us with a set of negative studies, without internal sensitivity checks, none of which is of adequate size to be totally convincing.

Opioids

Four opioid studies *(Table 55)* conform to the design required to answer the pre-emptive question. Katz and colleagues[13] looked at spinal opioid, while Richmond and colleagues,[16] Amanor-Boadu and colleagues[19] and Wilson and colleagues[22] investigated the intravenous route.

In the study by Katz and colleagues,[13] 30 thoracic patients were randomised to lumbar epidural fentanyl infusion, either pre- or intraoperatively *(Table 55)*. The infusion in the post-incision group was started 15 minutes after incision. Using VAS pain intensity and PCA intravenous morphine as outcome measures, they found significantly lower VAS pain intensity scores at 6 hours in the pre-operative group (with no significant difference in the PCA morphine consumption). From 12 hours to 24 hours, they found significantly lower PCA morphine consumption in the preoperative

TABLE 52 *Pre-emptive analgesia: studies included or excluded from the review*

NSAID	Local anaesthetic	Opioid
Studies included		
Flath, *et al.,* 1987[11]	Dahl, *et al.,* 1992[12]	Katz, *et al.,* 1992[13]
Sisk, *et al.,* 1989[14]	Pryle, *et al.,* 1993[15]	Richmond, *et al.,* 1993[16]
Sisk & Grover, 1990[17]	Dierking, *et al.,* 1992[18]	Amanor-Boadu, *et al.,* 1993[19]
Gustafsson, *et al.,* 1982[20]	Rice, *et al.,* 1990[21]	Wilson, *et al.,* 1994[22]
	Gunter, *et al.,* 1990[23]	
	Ejlersen, *et al.,* 1992[24]	
	Turner & Chalkiadis, 1994[25]	
NSAID + local anaesthetic + opioid: Kavanagh, *et al.,* 1994[26]		
Studies excluded		
Pre-emptive treatment with no postoperative comparison whether or not randomised		
Hutchison, *et al.,* 1990[27]	Tverskoy, *et al.,* 1990[28]	McQuay, *et al.,* 1988[29]
McGlew, *et al.,* 1991[30]	Jebeles, *et al.,* 1991[31]	Koskinen, *et al.,* 1991[32]
Smith & Brook, 1990[33]	Bugedo, *et al.,* 1990[34]	
Campbell, *et al.,* 1990[35]	Bach, *et al.,* 1988[36]	
	Narchi, *et al.,* 1991[37]	
RCT but pre-emptive treatment plus postoperative treatment compared with postoperative only		
Hill, *et al.,* 1987[38]		
Dupuis, *et al.,* 1988[39]		
Murphy & Medley, 1993[40]		

TABLE 53 *Pre-emptive analgesia: NSAID and paracetamol studies reviewed*

Study	Design	Number of patients	Procedure	Treatments	Outcome measures	Timing pre-operative	Timing post-operative	Outcome (pre- vs. post-operative)
Flath, *et al.,* 1987[11]	Parallel	120	Endodontic	Flurbiprofen, 100 mg, vs. placebo	categorical/ VAS PI	15 minutes	3 hours	NSD
Sisk, *et al.,* 1989[14]	Cross-over	20	Third molar	Diflunisal, 1 g, vs. placebo	categorical/ VAS PI	30 minutes	30 minutes	NSD
Sisk & Grover, 1990[17]	Cross-over	36	Third molar	Naproxen, 550 mg, vs. placebo	categorical/ VAS PI	30 minutes	30 minutes	NSD
Gustafsson, *et al.,* 1982[20]	Cross-over	50	Third molar	Paracetamol, 1 g, vs. placebo	VAS PI TFA	45 minutes	35 minutes	NSD

group with no significant difference in the VAS pain intensity score.

Richmond and colleagues[16] randomised 76 total abdominal hysterectomy patients to morphine, 10 mg intramuscular, as premedication, morphine, 10 mg intravenous, at induction or morphine, 10 mg intravenous, at closure *(Table 55)*. Analgesic outcome measures were VAS pain intensity scores and PCA intravenous morphine consumption. They found significantly lower PCA intravenous morphine consumption in the group who had received intravenous morphine at induction compared with the group who received the same dose by the same route at closure.

In a study of similar design, Amanor-Boadu and colleagues[19] looked at the effect of morphine, 5 mg intravenous, given at induction or at closure to 41 body surface surgery patients *(Table 55)*. Outcome measures were the time to first analgesic and the categorical and VAS pain intensity scores at that time. The categorical pain intensity scores were significantly lower at the time of remedication in the group given morphine at induction compared with the group given morphine at closure.

Wilson and colleagues[22] randomised 40 total abdominal hysterectomy patients to alfentanil, 40 μg/kg intravenous, at induction or after skin incision *(Table 55)*. Analgesic outcome measures

TABLE 54 *Pre-emptive analgesia: local anaesthetic studies reviewed*

Study	Design	No. of patients	Procedure	Treatments	Outcome measures	Timing pre-operative	Timing post-operative	Outcome (pre- vs. post-operative)
Epidural								
Dahl, *et al.*, 1992[12]	Parallel	32	Major colonic surgery	Epidural (T9–T12): bupivacaine and morphine	VAS and categorical PI at rest, cough and sitting up	bolus + infusion 40 minutes preoperative	bolus + infusion after surgery (still asleep)	NSD
Pryle, *et al.*, 1993[15]	Parallel	36	Abdominal hysterectomy	Epidural (lumbar): bupivacaine + adrenaline	VAS and categorical PI, PCA morphine, i.v.	40 minutes pre-incision	after surgery (still asleep)	NSD
Rice, *et al.*, 1990[21]	Parallel	40	Hernia, orchidopexy	Caudal: bupivacaine	Paediatric objective score, TFA	after induction	after surgery	NSD
Gunter, *et al.*, 1990[23]	Parallel	24	Distal hypospadias	Caudal: bupivacaine	TFA, analgesic needs	after induction	after surgery	NSD
Nerve block								
Dierking, *et al.*, 1992[18]	Parallel	32	Herniorr-rhaphy	Inguinal field block: lignocaine	VAS and categorical PI at rest, cough and moving	15 minutes preoperative	after closure	NSD
Infiltration								
Ejlersen, *et al.*, 1992[24]	Parallel	37	Hernior-rhaphy	Infiltration: lignocaine	TFA, analgesic needs	5 minutes pre-incision (19)	before closure (18)	Significantly delayed remedication time in preoperative group
Turner & Chalkiadis, 1994[25]	Parallel	90	Appendicec-tomy (29 controls with no infiltration)	Infiltration: lignocaine	VAS PI lying and sitting, PCA	3 minutes pre-incision (29)	at closure (32)	NSD on VAS PI or PCA

were VAS pain intensity scores and PCA intravenous morphine consumption. They found no significant difference in PCA morphine consumption but significantly higher pain scores at rest in the pre-emptive group.

Three studies suggest that opioids may have a pre-emptive effect. Unfortunately none of them is perfect. In the epidural study, there was a significant pre-emptive effect at only one of the six VAS pain intensity measurement points. In the morphine, 10 mg intravenous, study, PCA consumption was reversed in the subsequent 24 hours.[16] In the morphine, 5 mg intravenous, study only one of the two outcomes showed a significant effect. These three studies showing a weak positive pre-emptive effect with opioid are balanced by a negative effect.[22] The negative study is difficult to interpret because the difference in pain score at the same PCA consumption may mean a failure of study sensitivity.

Using pre-emptive NSAID, local anaesthetic and opioid together

Kavanagh and colleagues[26] compared a premedication of intramuscular morphine, 0.15 mg/kg,

TABLE 55 *Pre-emptive analgesia: opioid studies reviewed*

Study	Design	No. of patients	Procedure	Treatments	Outcome measures	Timing pre-operative	Timing post-operative	Outcome (pre- vs. post-operative)
Epidural								
Katz, *et al.*, 1992[13]	Parallel	30	Thoracotomy	Lumbar epidural: fentanyl	VAS PI and PCA morphine, i.v.	30-minute infusion started 55 minutes pre-incision	30-minute infusion started 15 minutes after incision	Significant difference (VAS PI lower at 6 hours and PCA morphine lower at 12–24 hours in pre-emptive group)
Intravenous								
Richmond, *et al.*, 1993[16]	Parallel	76	Total abdominal hysterectomy	Morphine, 10 mg, i.v. or i.m.	VAS PI and PCA morphine, i.v.	i.m. pre-medication (16) or i.v. at induction (23)	i.v. at closure (21)	Significant difference (PCA morphine lower to 24 hours in i.v. pre-emptive group at equivalent VAS PI)
Amanor-Boadu, *et al.*, 1993[19]	Parallel	41	Body surface	Morphine, 5 mg, i.v.	VAS and categorical PI at TFA	At induction (21)	At closure (20)	Significant difference (categorical PI lower at TFA in pre-emptive group)
Wilson, *et al.*, 1994[22]	Parallel	40	Total abdominal hysterectomy	Alfentanil, 40 µg/kg, i.v.	VAS PI and PCA morphine, i.v.	i.v. at induction (20)	i.v. 1 minute after incision (20)	Significant difference (VAS PI (rest) higher in pre-emptive group at same PCA)

perphenazine, 0.03 mg/kg, and rectal indomethacin, 100 mg, coupled with intercostal local anaesthetic with intramuscular midazolam premedication, 0.05 mg/kg, and saline intercostal blocks. Thirty thoracotomy patients were randomised to pre-emptive or control and compared using VAS pain intensity and PCA morphine consumption. There were no significant differences in VAS pain intensity scores. PCA consumption was significantly lower at 6 hours in the pre-emptive group, there was no significant difference at 12 hours and, at 24 hours and 48 hours, the morphine consumption in the pre-emptive group was significantly greater than control. This study did not compare the same intervention made before and after but has been included because it is the clearest example of an RCT using multiple (three drug classes) interventions to demonstrate a pre-emptive effect.

Comment

The evidence for pre-emptive effects should be answered separately for each of the three drug classes reviewed, because the answers may be different. For NSAIDs and paracetamol there are four good studies, all with no evidence of a pre-emptive effect. If there is a pre-emptive effect of NSAID it is unlikely to be seen with conventional dosing. In the case of local anaesthetics, one infiltration study showed a pre-emptive effect, another of similar design did not. Five other studies, spinal and nerve block, did not show any

effect. These negative local anaesthetic studies have been criticised[41] because they were 'contaminated' by opioid but, in two of the studies,[21,13] patients received no opioid. Perhaps a stronger criticism of these negative studies is that they lacked power.

The evidence with opioids is inconclusive. The evidence from fundamental studies for a pre-emptive effect of opioid[4,9] is stronger than its equivalents for local anaesthetic and NSAID. Inevitably what is now needed are studies of adequate design and size to establish whether or not there is indeed a measurable pre-emptive effect of opioid in man. If the intravenous route can be used to answer the question, studies are easier to perform than if the effect was only found with spinal routes. The caveat must be that the dose used in the (intrathecal) basic studies was large (up to 50 μg) and that the effect was demonstrated via the intrathecal route. This makes the human findings with relatively small intravenous doses all the more remarkable. Evidence of a pre-emptive effect with opioid would offer great potential benefit to patients with postoperative pain. It would also be important to know if such a pre-emptive effect applied to neuropathic pain.[9,36]

One important methodological issue is that increasingly investigators are using two post-operative outcomes, a VAS for pain intensity and PCA consumption. The assumption is made that patients will use PCA to achieve similar levels of VAS pain intensity. If the VAS pain intensity values for pre-emptive and control are not significantly different but the PCA consumption does show a significant difference, that is a valid result. The corollary is not valid. There may also be advantage (at least to the systematic reviewer) in using pain relief rather than pain intensity outcomes. Combining data across studies is much more valid for pain relief than pain intensity.[42] Ideally such combination of data would increase power and help to answer the clinical pre-emptive question.

One final point which gets forgotten is that acute tolerance is well known with opioids. Two of the pre-emptive studies[16,26] showed that pre-emptive treatment led to significant increase in postoperative analgesic consumption. It may be that any pre-emptive effect of opioids would be counteracted by induction of acute tolerance. This, however, is not in accord with the basic science demonstration of a pre-emptive opioid effect.

References

1. Woolf CJ. Evidence for a central component of post-injury pain hypersensitivity. *Nature* 1983;**306**:686–88.
2. Wall PD. The prevention of postoperative pain. *Pain* 1988;**33**:289–90.
3. McQuay HJ, Dickenson AH. Implications of nervous system plasticity for pain management. *Anaesthesia* 1990;**45**:101–2.
4. Dickenson AH, Sullivan AF. Subcutaneous formalin-induced activity of dorsal horn neurones in the rat: differential response to an intrathecal opiate administered pre or post formalin. *Pain* 1987;**30**:349–60.
5. Dickenson AH. Recent advances in the physiology and pharmacology of pain: plasticity and its implications for clinical analgesia. *J Psychopharmacol* 1991;**5**:342–51.
6. Coderre TJ, Vaccarino AL, Melzack R. Central nervous system plasticity in the tonic pain response to subcutaneous formalin injection. *Brain Res* 1990;**535**:155–8.
7. Malmberg AB, Yaksh TL. Hyperalgesia mediated by spinal glutamate of substance P receptor blocked by spinal cyclooxygenase inhibition. *Science* 1992;**257**:1276–9.
8. Selzer Z, Beilin BZ, Ginzburg R, Paran Y, Shimko T. The role of injury discharge in the induction of neuropathic pain behavior in rats. *Pain* 1991;**46**:327–36.
9. Puke MJC, Wiesenfeld-Hallin Z. The differential effects of morphine and the alpha2-adrenoreceptor agonists clonidine and dexmedetomidine on the prevention and treatment of experimental neuropathic pain. *Anesth Analg* 1993;**77**:104–9.
10. Jadad AR, McQuay HJ. A high-yield strategy to identify randomized controlled trials for systematic reviews. *Online J Curr Clin Trials* [serial online] 1993; Doc No 33:3973 words; 39 paragraphs; 5 tables.
11. Flath RK, Hicks ML, Dionne RA, Pelleu GB. Pain suppression after pulpectomy with preoperative flurbiprofen. *J Endodont* 1987;**13**:339–47.
12. Dahl JB, Hansen BL, Hjortsø NC, Erichsen CJ, Møiniche S, Kehlet H. Influence of timing on the effect of continuous extradural analgesia with bupivacaine and morphine after major abdominal surgery. *Br J Anaesth* 1992;**69**:4–12.
13. Katz J, Kavanagh BP, Sandler AN, Nierenberg H, Boylan JF, Friedlander M, *et al.* Preemptive analgesia: clinical evidence of neuroplasticity contributing to postoperative pain. *Anesthesiology* 1992;**77**:439–46.
14. Sisk AL, Mosley RO, Martin RP. Comparison of preoperative and postoperative diflunisal for suppression of postoperative pain. *J Oral Maxillofac Surg* 1989;**47**:464–8.

15. Pryle BJ, Vanner RG, Enriquez N, Reynolds F. Can pre-emptive lumbar epidural blockade reduce postoperative pain following lower abdominal surgery? *Anaesthesia* 1993;**48**:120–3.

16. Richmond CE, Bromley LM, Woolf CJ. Preoperative morphine pre-empts postoperative pain. *Lancet* 1993;**342**:73–5.

17. Sisk AL, Grover BJ. A comparison of preoperative and postoperative naproxen sodium for suppression of postoperative pain. *J Oral Maxillofac Surg* 1990;**48**:674–8.

18. Dierking GW, Dahl JB, Kanstrup J, Dahl A, Kehlet H. Effect of pre- vs postoperative inguinal field block on postoperative pain after herniorrhaphy. *Br J Anaesth* 1992;**68**:344–8.

19. Amanor-Boadu SD, Jadad AR, Glynn CJ, Jack TM, McQuay HJ. Influence of pre and postoperative morphine administration on pain after body surface surgery: a double-blind randomised controlled study. In: Proceedings 7th World Congress on Pain, 1993. Oxford: IASP/Elsevier;538–9.

20. Gustafsson I, Nystrom E, Quiding H. Effect of preoperative paracetamol on pain after oral surgery. *Eur J Clin Pharmacol* 1983;**24**:63–5.

21. Rice LJ, Pudimat MA, Hannalah RS. Timing of caudal block placement in relation to surgery does not affect duration of postoperative analgesia in paediatric ambulatory patients. *Can J Anaesth* 1990;**37**:429–31.

22. Wilson RJT, Leith S, Jackson IJB, Hunter D. Pre-emptive analgesia from intravenous administration of opioids. Anaesthesia 1994;49:591–3.

23. Gunter JB, Forestner JE, Manley CB. Caudal epidural anesthesia reduces blood loss during hypospadias repair. *J Urol* 1990;**144**:517–19.

24. Ejlersen E, Andersen HB, Eliasen K, Mogensen T. A comparison between preincisional and post-incisional lidocaine infiltration and postoperative pain. *Anesth Analg* 1992;**74**:495–8.

25. Turner GA, Chalkiadis G. Comparison of pre-operative with postoperative lignocaine infiltration on postoperative analgesia requirements. *Br J Anaesth* 1994;**72**:541–3.

26. Kavanagh BP, Katz J, Sandler AN, Nierenberg H, Roger S, Boylan JF, *et al.* Multimodal analgesia before thoracic surgery does not reduce postoperative pain. *Br J Anaesth* 1994;**73**:184–9.

27. Hutchison GL, Crofts SL, Gray IG. Preoperative piroxicam for postoperative analgesia in dental surgery. *Br J Anaesth* 1990;**65**:500–3.

28. Tverskoy M, Cozacov C, Ayache M, Bradley EL, Kissin I. Postoperative pain after inguinal herniorrhaphy with different types of anesthesia. *Anesth Analg* 1990;**70**:29–35.

29. McQuay HJ, Carroll D, Moore RA. Postoperative orthopaedic pain – the effect of opiate pre-medication and local anaesthetic blocks. Pain 1988;**33**:291–5.

30. McGlew IC, Angliss DB, Gee GJ, Rutherford A, Wood AT. A comparison of rectal indomethacin with placebo for pain relief following spinal surgery. *Anaesth Intensive Care* 1991;**19**:40–5.

31. Jebeles JA, Reilly JS, Gutierrez JF, Bradley EL, Kissin I. The effect of pre-incisional infiltration of tonsils with bupivacaine on the pain following tonsillectomy under general anesthesia. *Pain* 1991;**47**:305–8.

32. Koskinen R, Tigerstedt I, Tammisto T. The effect of peroperative alfentanil on the need for immediate postoeprative pain relief. *Acta Anaesthesiol Scand* 1991;**35** suppl 96:O21.

33. Smith AC, Brook IM. Inhibition of tissue prostaglandin synthesis during third molar surgery: use of preoperative fenbufen. *Br J Oral Maxillofac Surg* 1990;**44**:251–3.

34. Bugedo GJ, Cárcamo CR, Mertens RA, Dagnino JA, Muñoz HR. Preoperative percutaneous ilioinguinal and iliohypogastric nerve block with 0.5% bupivacaine for post-herniorrhaphy pain management in adults. *Reg Anesth* 1990;**15**:130–3.

35. Campbell WI, Kendrick R, Patterson C. Intravenous diclofenac sodium. Does its administration before operation suppress postoperative pain? *Anaesthesia* 1990;**45**:763–6.

36. Bach S, Noreng MF, Tjellden NU. Phantom limb pain in amputees during the first 12 months following limb amputation, after preoperative lumbar epidural blockade. *Pain* 1988;**33**:297–301.

37. Narchi P, Benhamou D, Fernandez H. Intra-peritoneal local anaesthetics for shoulder pain after day-case laparoscopy. *Lancet* 1991;**338**:291–5.

38. Hill CM, Carroll MJ, Giles AD, Pickvance N. Ibuprofen given pre and post operatively for the relief of pain. *Int J Oral Maxillofac Surg* 1987;**16**:420–4.

39. Dupuis R, Lemay H, Bushnelle MC, Duncan GH. Preoperative flurbiprofen in oral surgery: a method of choice in controlling postoperative pain. *Pharmacotherapy* 1988;**8**:193–200.

40. Murphy DF, Medley C. Preoperative indomethacin for pain relief after thoracotomy: comparison with postoperative indomethacin. *Br J Anaesth* 1993;**b**:298–300.

41. Katz J. Preop analgesia for acute pain. *Lancet* 1993;**342**:65–6.

42. Jadad-Bechara AR. Meta-analysis of randomised clinical trials in pain relief [DPhil thesis]. Oxford: University of Oxford, 1994.

Chapter 19

TENS in acute postoperative pain

Summary

The aim was to examine the evidence for the importance of randomisation of TENS in acute postoperative pain. Controlled trials were sought; randomisation and analgesic and adverse effect outcomes were summarised. A total of 46 reports were identified by searching strategies, of which 17 with 786 patients could be regarded unequivocally as RCTs in acute postoperative pain. No meta-analysis was possible. In 15 of the 17 RCTs, TENS was judged to have no benefit over placebo. Of the 29 excluded trials, 19 had pain outcomes but were not RCTs; in 17 of these, their authors concluded that TENS had a positive analgesic effect. No adverse effects were reported. Non-randomised trials overestimate treatment effects.

This chapter was published in full in 1996 by Carroll and colleagues.[1]

Introduction

TENS was originally developed as a way of controlling pain through the 'gate' theory.[2] There is conflicting professional opinion about the use of TENS in acute postoperative pain. The recommendations of the Agency for Health Care Policy and Research[3] for acute pain management state that TENS is "effective in reducing pain and improving physical function", while an earlier report from the UK College of Anaesthetists' working party on pain after surgery[4] says that "TENS is not effective as the sole treatment of moderate or severe pain after surgery". Some textbooks recommend or strongly recommend TENS for postoperative pain,[5–9] although one at least is uncertain.[10] TENS is of doubtful benefit in labour pain[11] and no systematic review of its use in chronic pain could be found.

Quality of methods used in clinical trials has been shown to be a key determinant of the eventual results. Schulz and colleagues[12] have demonstrated that trials which are not randomised or are inadequately randomised exaggerate the estimate of treatment effect by up to 40%. Studies which are not fully blinded can exaggerate the estimate of treatment effect by up to 17%. Evidence of the effect of randomisation in trials with pain as an outcome was sought in studies of TENS in acute postoperative pain.

Methods

A number of different search strategies were used to identify controlled trials for TENS in acute postoperative pain in both MEDLINE (1966–95: Knowledge Server v. 3.25: January 1996) and the Oxford Pain Relief Database (1950–92).[13] The terms 'TENS' and 'transcutaneous electrical nerve stimulation' were used in searching, including combinations of these words. Additional reports were identified from the reference lists of retrieved reports, review articles and textbooks.

Inclusion criteria were:

- full journal publication
- TENS
- postoperative pain with pain outcomes.

Reports of TENS for the relief of other acute pain conditions, such as labour pain, acute infections and procedures, or those in which there were less than ten patients per treatment group were excluded. Abstracts and review articles were not considered. Unpublished reports were not sought. Neither authors of reports nor manufacturers of TENS equipment were contacted.

Two types of control predominated: open studies compared TENS with conventional postoperative analgesia (intramuscular opiate) or with disabled TENS instruments (sham TENS). Some studies used blinded observers. While there was no prior hypothesis that TENS could not be blinded adequately, it was determined that, despite the considerable efforts documented in some reports, adequate blinding was impossible in practice.

Each report which could possibly meet the inclusion criteria was read by each author independently and scored for inclusion and quality using a 3-item scale.[14] Included reports had one point for randomisation, a further point if this had been done correctly, and a third if the number and reasons for withdrawals were given. Authors met to agree that studies were randomised, or whether

the description of the method of randomisation was adequate.[12]

Information about the surgery, numbers of patients, study design and duration of treatment was extracted from randomised reports. The type of TENS equipment, its settings and the method and frequency of its use and placement of electrodes was also extracted. Control group design and the use of TENS in these controls was similarly noted. Pain outcomes, overall findings and conclusions were noted for each report, together with any adverse effect information.

A judgement was then made as to whether the overall conclusion of randomised reports was positive or negative for the analgesic effectiveness of TENS. *Post-hoc* sub-group analysis in the original reports was not considered in the judgement of overall effectiveness. Reports which had pain measures but which were not randomised or were inadequately randomised were examined for positive or negative analgesic effectiveness of TENS using the judgement of their authors.

Results

Of the 46 reports that were considered, three did not have pain outcomes, three had fewer than ten patients per group, three had methodological problems and one reported on pain during rather than after a procedure. These were not considered further.

A total of 19 reports were either not RCTs or the method of randomisation was inappropriate *(Table 56)*.[15–33] Of the 19 reports with pain measures

TABLE 56 *Non-randomised reports of TENS in acute postoperative pain*

Study	Pain condition or operation type	Description	Authors' judgement about analgesic effectiveness
Ali, *et al.,* 1981[15]	Upper abdominal	Not RCT	Positive
Bussey & Jackson, 1981[16]	Cholecystectomy, hernia repair	Retrospective study not RCT	Positive
Cooperman, *et al.,* 1977[17]	Upper abdominal	Inadequate randomisation	Positive
Cornell, *et al.,* 1984[18]	Foot	Not RCT: matched case control	Positive
Hollinger, 1986[19]	Caesarian section	Not RCT	Positive
Hymes, *et al.,* 1974[20]	General	Not RCT	Positive
Issenman, *et al.,* 1985[21]	Spinal fusion	Not RCT	Positive
Lanham, *et al.,* 1984[22]	Foot	Retrospective, not RCT	Positive
Merrill, 1987[23]	Urological	Not RCT	Positive
Merrill, 1988[24]	Urological	Not RCT	Positive
Merrill, 1988[25]	Urological	Not RCT	Positive
Neary, 1981[26]	Abdominal, thoracic	Not RCT	Positive
Reuss, *et al.,* 1988[27]	Cholecystectomy	Not RCT	Negative
Rooney, *et al.,* 1983[28]	Thoracotomy	Inadequate randomisation method	Positive
Schomberg & Carter-Baker, 1983[29]	Laparotomy	Retrospective, not RCT	Positive
Schuster & Infante, 1980[30]	Low back	Not RCT	Positive
Solomon, *et al.,* 1980[31]	Lumbar, hip, gynaecological	Retrospective, not RCT	Positive
Stabile & Mallory, 1978[32]	Knee and hip joint	Not RCT	Positive
Strayhorn, 1983[33]	Gastric bypass	Not RCT	Negative

excluded because they were either not randomised or inadequately randomised, 17 were judged by their authors to have positive analgesic results for TENS in acute postoperative pain.

Of the 17 randomised studies with pain outcomes found, 15 were judged to show no analgesic benefit of TENS in acute postoperative pain *(Table 57)*.

TABLE 57 *Analgesic benefit of TENS in postoperative pain*

	Analgesic result	
	Positive	**Negative**
Randomised	2	15
Inadequate or not randomised	17	2

Randomised studies

The randomised studies had information from 786 patients *(Table 58)*.[27,34–49] TENS was used after various operative procedures including cardiothoracic, major orthopaedic and gastrointestinal surgery. Ten different TENS machines were used with different control settings and durations of treatment; in six studies, individual titration of settings was reported. TENS was compared with sham TENS without batteries, with batteries reversed or with sub-threshold stimulation in 14 studies; in the other three, TENS plus intramuscular opiate was compared with intramuscular opiate alone. Quality scores were generally 1 or 2 out of a maximum of 3. The most common outcome measures reported were analgesic consumption and a variety of pain score measurements. Information was not presented in formats which allowed extraction for meta-analysis *(Table 58)*.

TENS versus sham TENS

Of the 17 included RCTs, 14 compared TENS with sham TENS. Not one found any difference. One of the 14[49] reported no significant difference between TENS and sham TENS for analgesic consumption but did report a statistically significant difference for pain intensity in favour of the active TENS; the published results, however, used a one-tailed statistical test which was judged inappropriate.

TENS versus opiate control

Of the 17 included RCTs, seven compared opiate plus TENS with opiate alone, four of which also included sham TENS. Of the seven studies, five failed to detect any differences in analgesic consumption or pain measurements between TENS and non-TENS controls. Two reports were judged by their authors and by us to be positive.[43,47]

Pike[43] studied 40 patients after total hip replacement. The study had as its main outcome measure the number of pethidine (meperidine) injections in the first 2 postoperative days and a retrospective global rating. Patients with active TENS had significantly fewer pethidine injections on the first postoperative day as well as higher scores on global rating of treatment. VanderArk and McGrath[47] recruited 100 patients having abdominal and thoracic surgery in 2 months and, although there was more success with active TENS used for 20 minutes three times a day, maximal relief was "almost invariably associated with the first stimulation". Generally there were no obvious differences between the use of TENS in these two positive studies and the 15 which showed no benefit.

Adverse events

No report described systematic recording of adverse events nor were any reported.

Comment

The gold standard in clinical trials is adequate randomisation.[12] For nearly 20 years, non-randomised studies have been shown to yield larger estimates of treatment effects than studies using random allocation.[50] The degree of the exaggeration of treatment effect when randomisation is inappropriate can be as much as 40%.[12] These findings underpin the inclusion criteria chosen in systematic reviews.

For TENS in acute postoperative pain, 17 of 19 reports with pain outcomes which were either not randomised or inappropriately randomised claimed TENS to be effective, compared with two of 17 RCTs *(Table 57)*.

The possibility of bias exists. The method of randomisation was described in only two reports.[17,28] The method described was inadequate in both, one using a nurse to randomise patients[17] and the other using alternate allocation.[28] Reports which said only that they were randomised may also have used an inadequate method.

That these data represent the lowest common denominator of information, essentially vote counting rather than a more sophisticated analysis, reflects the nature of the analgesic scoring methods that predominated in the original reports. Pain scoring using analogue or categorical scales was reported as a mean (an unreliable statistic[51]); alternatively, mean analgesic consumption or time to first analgesic was used. None of these allowed data extraction for further statistical analysis or

TABLE 58 *Randomised studies of TENS in acute postoperative pain*

Study	Operation type	Study design and duration of treatment periods	Number of patients	TENS details	TENS control setting	TENS control	Pain outcomes	Results for pain outcomes	Judgement	Score
Conn, *et al.*, 1986[34]	Appendicectomy	Parallel group: TENS 15; sham TENS 13; stand ard postoperative analgesia 14. 48 hours	42	Dow Corning Wright, single channel, electrodes (either side of wound).	Fixed rate (tingling sensation preoperatively).	Sham TENS (not turned on)	VAS PI at 48 hours; analgesic consumption (24, 48 hours).	NSD between sham and active TENS for pain and drug consumption; significant difference for PI control vs. TENS and sham TENS ($p < 0.01$).	Negative	1
Cushieri, *et al.*, 1985[35]	Abdominal	Parallel group: sham TENS 53; TENS 53. 72 hours	106	Codman, dual channel, 2 electrodes (either side of wound).	Fixed rate (tingling sensation pre-operatively), rectangular wave form, pulse width (170 ms), pulse rate 80/second, output 15 milliamps.	Sham TENS (batteries reversed)	VAS PI: average pain twice daily, morphine consumption.	NSD between TENS and sham TENS.	Negative	2
Davies, 1982[36]	Caesarean section	Parallel group (female): general anaesthetic + TENS 10; epidural + TENS 11; general anaes-thetic + sham TENS 8; epidural + sham TENS 6. 24 hours	35	Stim-Tec EPC Mini, Model 6011, dual channel (1 channel only used), 2 electrodes (either end of wound).	Amplitude individually titrated, wave fixed (during surgery).	Sham TENS (no batteries)	VAS PI: hourly, time to first analgesic, analgesic consumption.	No overall difference in analgesic consumption or pain.	Negative	2
Forster, *et al.*, 1994[37]	Coronary artery bypass	Parallel group (males): TENS 15; sham TENS 15; postoperative analgesia 15. 72 hours	45	Nuwave, Staodyn. 1 pair electrodes (T1–T5), 1 pair (either side of wound).	Individually titrated (tingling sensation).	Sham TENS (no current)	Pain (0–10) on cough and rest; narcotic consumption.	NSD TENS vs. sham TENS.	Negative	2
Galloway, *et al.*, 1984[38]	Cholecystectomy	Parallel group: TENS 14; remote TENS 12; post-operative analgesia 14. 48 hours	40	3M Tenzcare, dual channel, site of electrodes not described.	Individually titrated.	Sham TENS (remote non-segmental)	VAS PI; categorical PI (4-point scale) at 24, 48 hours; analgesic consumption.	NSD TENS vs. sham TENS for pain or analgesic consumption. Pain but not analgesic consumption significantly worse in control group ($p < 0.05$) immediately after surgery.	Negative	1
Gilbert, *et al.*, 1986[39]	Herniorrhaphy	Parallel group (males): TENS 20; sham TENS 20. 72 hours	40	Dow Corning, Wright Care, dual channel, 2 electrodes (either side of wound).	Individually titrated: pulse duration 180 µs, frequency 70 Hz, amplitude 7.5.	Sham TENS (no current)	VAS PI twice daily; analgesic consumption.	NSD TENS vs. sham TENS.	Negative	1
Lim, *et al.*, 1983[40]	Abdominal	Parallel group: TENS 15; sham TENS 15. 48 hours	34	Neuromed 3722, 2 electrodes (either side of wound).	Individually titrated (tingling sensation).	Sham TENS (batteries reversed)	VAS PI (2, 4, 6, 24, 48 hours); analgesic consumption.	NSD TENS vs. sham TENS.	Negative	2
McCallum, *et al.*, 1988[41]	Laminectomy	Parallel group: TENS 10; sham TENS 10. 24 hours	20	Dow Corning, Wright Care, dual channel, 4 electrodes (at each end and on either side of wound).	Individually titrated, 180 µs pulse width, frequency 70 Hz.	Sham TENS (no current)	PCA morphine consumption, 24 hours.	NSD TENS vs. sham TENS.	Negative	2
Navarathnam, *et al.*, 1984[42]	Cardiac	Parallel group: TENS 14; sham TENS 17. 72 hours	31	3M Tenz care Model 6240, dual channel, 2 pairs electrodes (either side of the wound and mid-thoracic region).	Individually titrated, pulse rate 5, width control 3, amplitude.	Sham TENS (no batteries)	5-point categorical PI; analgesic consumption.	NSD TENS and sham TENS.	Negative	2

continued

TABLE 58 contd *Randomised studies of TENS in acute postoperative pain*

Study	Operation type	Study design and duration of treatment periods	Number of patients	TENS details	TENS control setting	TENS control	Pain outcomes	Results for pain outcomes	Judgement	Score
Pike, 1978[43]	Total hip replacement	Parallel group: TENS 20; opiate control 20. 24 hours	40	EPC TimeTech clinical stimulator, dual channel, 2 pairs electrodes (1 pair paravertebrally (L2–S2), between trochanter and coccyx, 1 pair above iliac crest, head of fibula).	Individually titrated, continual stimulation.		Global assessment; analgesic consumption.	Significantly less pethidine consumed in TENS group on day 1 ($p < 0.001$).	Positive	1
Reuss, *et al.*, 1988[27]	Cholecystectomy	Parallel group: TENS 30; opiate control 34.	64	EPC, electrodes placed within 2 cm of the wound.	Pulse rate 50/second, pulse width 170 ms, amplitude 0–50.		Daily dose of pethidine for 3 postoperative days.	NSD	Negative	1
Smedley, *et al.*, 1988[44]	Inguinal hernia repair	Parallel group (males): TENS 34; sham TENS 28. 48 hours	62	3M Tenzcare dual channel, 2 pairs of electrodes (over first lumbar vertebrae and on either side of wound).	Individually titrated (tingling sensation), 70 Hz rectangular pulse, amplitude.	Sham TENS (controls turned off)	VAS PI, 6, 12, 24, 36, 48 hours; opiate consumption.	NSD TENS vs. sham TENS.	Negative	2
Stubbing & Jellicoe, 1988[45]	Thoracotomy	Parallel group: TENS + i.m. omnopon 20; i.m. omnopon 20. 48 hours	40	Dow Corning, Wright Care 2 channel, 2 electrodes (either side of incision).	Individually titrated, fixed pulse rate 70/second, rectangular waveform, pulse width 180 µs.		5-point PI, 6, 24, 48 hours; analgesic consumption; time to oral analgesics; length of hospital stay.	NSD TENS vs. sham TENS.	Negative	1
Taylor, *et al.*, 1983[46]	Abdominal	Parallel group: TENS 30; sham TENS 22; i.m. narcotics 25. 72 hours	77	MedGen, electrode, placement not described.	Fixed pulse width 80 ms, frequency 40 Hz, amplitude individually titrated.	Sham TENS (no current)	Daily 10-point PI.	NSD TENS vs. sham TENS or control.	Negative	1
VanderArk & McGrath, 1975[47]	Abdominal and thoracic	Parallel group: TENS 61; sham TENS 39. 24 hours post-surgery until discharge, TDS x 20 minutes.	100	Neuromed Model 3700, Meditronic, electrode site individually chosen.	Frequency 100–150/second, output 20–35, pulse duration 250–400 ms.	Sham TENS (no batteries)	Pain; analgesic consumption; duration of relief.	Significant difference reported. 2/39 partial relief or complete relief sham TENS vs. 34 /61 with active TENS. Analgesic consumption not reported.	Positive	1
Walker, *et al.*, 1991[48]	Total knee replacement	Phase 2 – parallel group: TENS 18; sham TENS 18; post-operative analgesia 12. 72 hours	48	Strodynamics, continuous. No other information given, electrode placement not described.	Amplitude setting individually titrated, pulse duration 100 µs at 70/second.	Sham TENS (sub-threshold stimulation)	Analgesic consumption; length of hospital stay.	NSD TENS vs. sham TENS or control.	Negative	1
Warfield, *et al.*, 1985[49]	Thoracotomy	Parallel group: TENS 12; sham TENS 12. 48 hours	24	3M Tenzcare 6240, 2 electrodes placed on either side of incision.	Continuous stimulation, amplitude 7, pulse rate 3, pulse width 5.	Sham TENS (no current)	PI 0–10; analgesic consumption.	NSD TENS vs. sham TENS. Positive result reported with one-tailed test of statistical significance.	Negative	2

comparison between reports. While more rigorous pain scoring might have been used, there is no evidence that all of the reports suffered a systematic failure in analgesic measurement.

Inadequacy of blinding in clinical trials of anal-gesic interventions continues to be of concern,[52] although this may be less of an issue with pharmacological interventions.[51] Blinding of procedures is much more difficult than blinding of drug studies. Most of the TENS studies did make attempts at blinding, for instance by removing batteries from the TENS apparatus (sham TENS) or by using staff with no knowledge of the study or allocation to conduct the

patient assessments. Lack of blinding has been estimated to exaggerate the estimate of treatment effect of trials by some 17%.[12] Adequate blinding of TENS for both carers and patients is particularly difficult.[53] None of the reports was judged to have been blinded and this lowered the quality scores given to the 17 randomised studies. The fact that only two of the reports showed any positive effect of TENS in acute postoperative pain is all the more striking because of this potential overestimation of treatment effect due to lack of blinding.

The clear message from the studies considered in this systematic review is that adequate randomisation is an important quality standard in studies with pain outcomes. Including non-randomised studies in reviews may give the wrong answer. The Agency for Health Care Policy and Research guidelines on acute pain management included non-randomised reports, and this may explain their more positive attitude towards TENS.[3]

References

1. Carroll D, Tramèr M, McQuay H, Nye B, Moore A. Randomization is important in studies with pain outcomes: systematic review of transcutaneous electrical nerve stimulation in acute postoperative pain. *Br J Anaesth* 1996;**77**:798–803.

2. Melzack R, Wall PD. Pain mechanisms: a new theory. *Science* 1965;**150**:971–8.

3. Acute Pain Management Guideline Panel. Acute pain management: operative or medical procedures and trauma. Clinical Practice Guideline No. 1. Agency for Health Care Policy and Research, US Department of Health and Human Services. AHCPR Publication No. 92-0032. Rockville, MD: Public Health Service, 1992:24–5.

4. Royal College of Surgeons of England and the College of Anaesthetists. Report of the working party on pain after surgery. London: Royal College of Surgeons, 1990.

5. Charman RA. Physiotherapy for the relief of pain. In: Carroll D, Bowsher D, editors. Pain management and nursing care. Oxford: Butterworth-Heinemann, 1993:146–65.

6. Woolf CJ, Thomson JW. Stimulation-induced analgesia: transcutaneous electrical nerve stimulation (TENS) and vibration. In: Wall PD, Melzack R, editors. Textbook of pain. 3rd edition. Edinburgh: Churchill Livingstone, 1994:1191–217.

7. McCaffery M, Beebe A. Pain: a clinical manual for nursing practice. St. Louis, Missouri: CV Mosby Company, 1989.

8. Chaney ES. The management of postoperative pain. In: Wells P, Frampton V, Bowsher D, editors. Pain: management and control in physiotherapy. Oxford: Heinemann, 1988:253–68.

9. Park G, Fulton B. The management of acute pain. Oxford: Oxford University Press, 1991.

10. Justins DM, Richardson PH. Clinical management of acute pain. *Br Med Bull* 1991;**47**:561–83.

11. Carroll D, Tramèr M, McQuay H, Nye B, Moore A. Transcutaneous electrical nerve stimulation in labour pain: a systematic review. *Br J Obstet Gynaecol* 1997;**104**:169–75.

12. Schulz KF, Chalmers I, Hayes RJ, Altman DG. Empirical evidence of bias: dimensions of methodological quality associated with estimates of treatment effects in controlled trials. *JAMA* 1995;**273**:408–12.

13. Jadad AR, McQuay HJ. A high-yield strategy to identify randomized controlled trials for systematic reviews. *Online J Curr Clin Trials* [serial online] 1993; Doc No 33:3973 words; 39 paragraphs; 5 tables.

14. Jadad AR, Moore RA, Carroll D, Jenkinson C, Reynolds DJM, Gavaghan DJ, *et al.* Assessing the quality of reports of randomized clinical trials: is blinding necessary? *Control Clin Trials* 1996;**17**:1–12.

15. Ali J, Yaffe CS, Serrette C. The effect of transcutaneous electric nerve stimulation on postoperative pain and pulmonary function. *Surgery* 1981;**89**:507–12.

16. Bussey JG, Jackson A. TENS for postsurgical analgesia. *Contemp Surg* 1981;**18**:35–41.

17. Cooperman AM, Hall B, Mikalacki K, Hardy R, Sadar E. Use of transcutaneous electrical stimulation in the control of postoperative pain. *Am J Surg* 1977;**133**:185–7.

18. Cornell PE, Lopez AL, Malofsky H. Pain reduction with transcutaneous electrical nerve stimulation after foot surgery. *J Foot Surg* 1984;**23**:326–33.

19. Hollinger JL. Transcutaneous electrical nerve stimulation after cesarean birth. *Phys Ther* 1986;**66**:36–8.

20. Hymes AC, Raab DE, Yonehiro EG, Nelson GD, Printy AL. Electrical surface stimulation for control of acute postoperative pain and prevention of ileus. *Surg Forum* 1974;**24**:447–9.

21. Issenman J, Nolan MF, Rowley J, Hobby R. Transcutaneous electrical nerve stimulation for pain control after spinal fusion with Harrington rods. *Phys Ther* 1985;**65**:1517.

22. Lanham RH, Powell S, Hendrix BE. Efficacy of hypothermia and transcutaneous electrical nerve stimulation in podiatric surgery. *J Foot Surg* 1984;**23**:152–8.

23. Merrill DC. Electroanalgesia in urologic surgery. *Urology* 1987;**29**:494–7.

24. Merrill DC. FasTENS – a disposable transcutaneous electrical nerve stimulator designed specifically for use in postoperative patients. *Urology* 1988;**31**:78–9.

25. Merrill DC. Clinical evaluation of FasTENS, an inexpensive, disposable transcutaneous electrical nerve stimulator designed specifically for post-operative electroanalgesia. *Urology* 1989;**33**:27–30.

26. Neary JM. Transcutaneous electrical nerve stimulation for the relief of post-incisional surgical pain. *Am Assoc Nurs J* 1981;**49**:151–5.

27. Reuss R, Cronen P, Abplanalp L. Transcutaneous electrical nerve stimulation for pain control after cholecystectomy: lack of expected benefits. *South Med J* 1988;**81**:1361–3.

28. Rooney SM, Jain S, Goldiner PL. Effect of transcutaneous nerve stimulation on postoperative pain after thoracotomy. *Anesth Analg* 1983;**62**:1010–12.

29. Schomberg FL, Carter-Baker SA. Transcutaneous electrical nerve stimulation for postlaparotomy pain. *Phys Ther* 1983;**63**:188–93.

30. Schuster GD, Infante MC. Pain relief after low back surgery: the efficacy of transcutaneous electrical nerve stimulation. *Pain* 1980;**8**:299–302.

31. Solomon R, Viernstein MC, Long D. Reduction in postoperative pain and narcotic use with transcutaneous electrical nerve stimulation. *Surgery* 1980;**87**:142–6.

32. Stabile ML, Mallory TH. The management of postoperative pain in total joint replacement. *Orthop Rev* 1978;**7**:121–3.

33. Strayhorn G. Transcutaneous electrical nerve stimulation and postoperative use of narcotic analgesics. *J Natl Med Assoc* 1983;**75**:811–16.

34. Conn IG, Marshall AH, Yadav SN, Daly JC, Jaffer M. Transcutaneous electrical nerve stimulation following appendicectomy: the placebo effect. *Ann R Coll Surg Engl* 1986;**68**:191–2.

35. Cuschieri RJ, Morran CG, McArdle CS. Transcutaneous electrical stimulation for postoperative pain. *Ann R Coll Surg Engl* 1985;**67**:127–9.

36. Davies JR. Ineffective transcutaneous nerve stimulation following epidural anaesthesia. *Anaesthesia* 1982;**37**:453–7.

37. Forster EL, Kramer JF, Lucy SD, Scudds RA, Novick RJ. Effect of TENS on pain, medications, and pulmonary function following coronary artery bypass graft surgery. *Chest* 1994;**106**:1343–8.

38. Galloway DJ, Boyle P, Burns HJ, Davidson PM, George WD. A clinical assessment of electro-analgesia following abdominal operations. *Surg Gynecol Obstet* 1984;**159**:453–6.

39. Gilbert JM, Gledhill T, Law N, George C. Controlled trial of transcutaneous electrical nerve stimulation (TENS) for postoperative pain relief following inguinal herniorrhaphy. *Br J Surg* 1986;**73**:749–51.

40. Lim AT, Edis G, Kranz H, Mendelson G, Selwood T, Scott DF. Postoperative pain control: contribution of psychological factors and transcutaneous electrical stimulation. *Pain* 1983;**17**:179–88.

41. McCallum MI, Glynn CJ, Moore RA, Lammer P, Phillips AM. Transcutaneous electrical nerve stimulation in the management of acute postoperative pain. *Br J Anaesth* 1988;**61**:308–12.

42. Navarathnam RG, Wang IY, Thomas D, Klineberg PL. Evaluation of the transcutaneous electrical nerve stimulator for postoperative analgesia following cardiac surgery. *Anaesth Intensive Care* 1984;**12**:345–50.

43. Pike PM. Transcutaneous electrical stimulation. Its use in the management of postoperative pain. *Anaesthesia* 1978;**33**:165–71.

44. Smedley F, Taube M, Wastell C. Transcutaneous electrical nerve stimulation for pain relief following inguinal hernia repair: a controlled trial. *Eur Surg Res* 1988;**20**:233–7.

45. Stubbing JF, Jellicoe JA. Transcutaneous electrical nerve stimulation after thoracotomy. Pain relief and peak expiratory flow rate: a trial of transcutaneous electrical nerve stimulation. *Anaesthesia* 1988;**43**:296–8.

46. Taylor AG, West BA, Simon B, Skelton J, Rowlingson JC. How effective is TENS for acute pain? *Am J Nurs* 1983;**83**:1171–4.

47. VanderArk GD, McGrath KA. Transcutaneous electrical stimulation in treatment of postoperative pain. *Am J Surg* 1975;**130**:130–8.

48. Walker RH, Morris BA, Angulo DL, Schneider J, Colwell CWJ. Postoperative use of continuous passive motion, transcutaneous electrical nerve stimulation, and continuous cooling pad following total knee arthroplasty. *J Arthroplasty* 1991;**6**:151–6.

49. Warfield CA, Stein JM, Frank HA. The effect of transcutaneous electrical nerve stimulation on pain after thoracotomy. *Ann Thorac Surg* 1985;**39**:462–5.

50. Chalmers TC, Matta RJ, Smith JJ, Kunzler AM. Evidence favoring the use of anticoagulants in the hospital phase of acute myocardial infarction. *N Engl J Med* 1977;**297**:1091–6.

51. McQuay H, Carroll D, Moore A. Variation in the placebo effect in randomised controlled trials of analgesics: all is as blind as it seems. *Pain* 1996;**64**:331–5.

52. Wall PD. The placebo effect: an unpopular topic. *Pain* 1992;**51**:1–3.

53. Deyo RA, Walsh NE, Schoenfeld LS, Ramamurthy S. Can trials of physical treatments be blinded? The example of transcutaneous electrical nerve stimulation for chronic pain. *Am J Phys Med Rehabil* 1990;**69**:6–10.

Chapter 20

TENS in labour pain

Summary

TENS is used widely for relief of pain in labour. However, two previous systematic reviews have questioned its effectiveness in this context. Reports were sought by searching MEDLINE, EMBASE, CINAHL and the Oxford Pain Relief Database. Outcomes included pain and adverse effect measures.

Ten RCTs were found; of the 877 women involved, 436 received active TENS and 441 acted as controls (sham TENS or no treatment). There were no significant differences reported for prospective primary pain outcomes in any of the ten studies. Three studies reported significant differences between active and sham TENS for secondary pain outcomes. The use of additional analgesic interventions was not different with active or sham TENS (relative risk 0.88 (095% CI, 72–1.07)).

The findings suggest that TENS has no significant effect on pain in labour. Women in labour should be offered more effective interventions for the relief of pain.

Methods

A number of different search strategies were used to identify eligible reports of RCTs of TENS in labour pain in MEDLINE (1966–97), EMBASE (1980–97), CINAHL (1982–97), the Cochrane Library (issue 2, 1997) and the Oxford Pain Relief Database (1950–95).[1] The date of last search was April 1997. The terms 'TENS', 'transcutaneous electrical nerve stimulation', 'labour' and 'childbirth' were used in searching, including in combinations, and there was no language restriction. Additional reports were identified from the reference lists of retrieved reports, review articles[2] and textbooks. Manufacturers of TENS equipment were not contacted. Abstracts and review articles were not considered. Unpublished reports were not sought.

Inclusion criteria were full journal publication, TENS, labour pain with pain outcomes and randomised treatment allocation. Reports of TENS for the relief of other pain conditions or those in which there were fewer than ten patients per treatment group were excluded.[3]

Each report which could possibly meet the inclusion criteria was read by each author independently and scored for inclusion and quality using a 3-item scale[4] which examined randomisation, blinding and withdrawal, and drop-outs. An included report could have a maximum score of 5 and a minimum of 1. Where the method of treatment allocation was unconcealed (alternate allocation, for instance) the report was excluded. A *pre-hoc* judgement was made that it would be difficult to blind TENS and thus quality scores were unlikely to exceed 3.

Information about inclusion criteria for women in labour, stage of labour, cervical dilatation, number of women, study design and timing and duration of treatment was extracted from the reports, together with information on other analgesic interventions and preferences for future childbirth. The type of TENS equipment, its settings and the method and frequency of its use and positioning of electrodes was also extracted. Control group design and the use of TENS in these controls was similarly noted, including the methods used to disable TENS devices (e.g. sham TENS with no battery).

The effectiveness of TENS was judged by whether or not a statistically significant difference between TENS and the control group (sham TENS or no treatment) was reported in the original report for at least one of the outcome measures used. Outcomes were judged by us as being either primary or secondary. Primary outcomes were defined as any prospective assessment of pain intensity or relief made at the time of labour and when TENS was in use. Secondary outcomes were defined as any retrospective assessment of pain or pain relief or any other measure, or judgement made after delivery, or after TENS had been discontinued. Secondary outcomes included the use of any additional pain interventions, the timing of such interventions and any retrospective global evaluation of the study treatments. A judgement was then made as to whether the overall conclusion of the report was positive or negative for the anal-

gesic effectiveness of TENS on primary and secondary outcomes separately. *Post-hoc* sub-group analysis in the original reports was not considered in the judgement of overall effectiveness. Any information on adverse effects was summarised.

Relative risk or benefit was calculated with the 95% CI using a random effect model[5] for analgesic data which were not homogeneous ($p < 0.1$). A statistically significant difference from control was assumed when the 95% CI of relative risk did not include 1. An NNT was calculated with 95% CIs[6] for any comparison which showed significance with relative risk.

Results

Two additional reports were found which were not included in a previous review.[7] Ten reports involving 877 women were included; 436 women received active TENS and 441 acted as controls. The methodological details of the study designs, instructions to women before and during labour, TENS details and settings, control conditions and methods of blinding are presented in *Table 59*. One study[8] used cranial TENS; others used TENS with dorsal or suprapubic stimulation. Nine different TENS devices were used in the ten studies, predominantly with individual titration.

Three studies used conventional analgesic administration (no TENS) as the control group.[8–10] In seven studies disabled TENS instruments (sham TENS) were used as a control group.[11–17] In one study both sham TENS and a no TENS control were used.[13] In only one study[16] was a sufficiently determined attempt at blinding made to merit any inclusion points for blinding. This study had an inclusion quality score of 4; seven studies scored 2 and two scored 1.

Pain outcomes and results for the ten studies are presented in *Table 60*. There was no consistency in the method of measuring pain intensity or relief. In some studies suprapubic and back pain was measured separately, and in others pain was measured at different stages of labour or at different degrees of cervical dilatation. No study recorded any difference in pain intensity or relief scores between TENS and control during labour.

Additional analgesic interventions were recorded in eight of the ten reports.[9,11–17] In two studies, the total number of interventions was noted. Bundsen[9] recorded 17 additional analgesic interventions in 11 women receiving usual obstetric analgesic care compared with 21 additional interventions in 16 women receiving TENS. Nesheim[14] recorded that 35 women with TENS needed 49 analgesic interventions compared with 63 interventions in 35 women with sham TENS. In one study,[13] all analgesic interventions were reported, other than epidurals which were discouraged in this study.

Figures for the number of women who received any other analgesic intervention were given in five studies.[11–13,15,16] The results of this secondary outcome for the comparisons of active TENS with sham TENS, together with the number of women in the comparison, are presented in *Figure 47*. Overall, of 292 women, 227 (78%) given active TENS had an additional analgesic intervention compared with 239 of 280 women (85%) having sham TENS. There was no difference between active and sham TENS in the three largest studies *(Figure 47)*. The combined result of all five studies had a relative risk of 0.88 (95% CI, 0.72–1.07; *Figure 48)*. The lack of any statistical difference made the calculation of an NNT irrelevant.

None of the studies were judged to have a positive result for the primary outcome measures, which were prospective measures of pain intensity or relief. For the secondary outcomes of additional analgesics taken or time to next analgesic, three studies[8,11,15] were judged to have a positive result. The three positive studies included a study of cranial TENS in 20 women,[16] in which only the single outcome of other analgesics taken was used, and a comparison of TENS plus epidural compared with epidural alone.[17]

There were no reports of adverse events in any of the ten studies.

Comment

None of the ten studies included in this review reported any significant difference between the active TENS treatments and controls for any of the primary pain outcome measures used. This strengthens the findings of previous negative reviews on TENS in labour.[2,18] The weak evidence from secondary outcome measures that the need for additional analgesics may be diminished[18] was negated by an additional large trial,[13] in which no difference was found between active TENS and sham TENS. Trial size is likely to be important when assessing even primary outcomes but for weak secondary pain outcomes, such as additional analgesic requirements, the effect of trial size shown in *Figure 47* was dramatic. The two trials

TABLE 59 *Randomised studies of TENS in labour pain*

Study	Early criteria	Study design and duration of treatment periods	Pre-study instructions	Intra-study instructions	TENS machine	TENS setting	Electrode details	Control group(s)	Blinding	Quality score
Bunsden, *et al.*, 1982[9]	Induced labour only (amniotomy ± oxytocin), who did not desire specific alternative pain intervention, all attended antenatal clinic pre-delivery.	Parallel group: 1. TENS (16); 2. control group (11). Dorsal and suprapubic stimulation at two different frequencies; TENS from time of first contraction to parturition.	Standard information about study and available methods of PR. Women requested to try TENS before receiving other pain interventions.	Women in control group allowed conventional obstetric analgesia as required.	Custom-built stimulator	Individually titrated by one author: two frequencies (high frequency 50 Hz, pulse train), until paraesthesia to painful areas.	Two electrodes: one supra-pubic, one low back.	Individually titrated conventional obstetric analgesia (nitrous oxide, epidural, pethidine, diazepam, pudendal block, paracervical block).	Open, no attempt to blind. TENS given by one of authors over painful areas (low back or supra-pubic) randomly; if first method not effective or caused discomfort then other site used after 15–30 minutes.	2
Champagne, *et al.*, 1984[11]	Primipara and multipara requiring analgesia during labour and delivery.	Parallel group: 1. Limoge cranio TENS (10); 2. Sham cranio TENS (10). Began when analgesia requested.	Not described.	Not described.	Anesthelec MPO2 cranial Limoge TENS	High frequency, 166 kHz, 1.2 ms, 20%; low frequency, 83 Hz, 4 ms, 33%.	Three electrodes: two post mastoid, one between eyebrows.	Sham cranio TENS; (no current).	Described as double-blind, light hidden on both machines, independent person set up machine. Patient and observer blind.	2
Chia, *et al.*, 1990[10]	Surgical induction or early labour, primigravida who had not previously experienced other forms of analgesia in labour.	Partial crossover: 1. TENS (10); 2. Entonox (10). Study began when patients requested analgesia in early stage of labour; complete once second stage of labour reached or other form of analgesia requested.	TENS unit explained and demonstrated before painful contractions, midwife taught patients breathing technique for Entonox.	Could use burst mode if needed during painful contractions.	Spembly (Obstetric Pulsar) dual channel	Individually titrated, fixed pulse rate of 200 ms, amplitude 48 mA	Two pairs, one either side of midline 5 cm apart; one pair at T11, T10; 1 pair upper sacral vertebrae.	Entonox.	Open; no attempt to blind treatments.	2
Harrison, *et al.*, 1986[12]	Primigravida and third labour, who did not desire specific alternative pain intervention.	Parallel group: 1. Parity 0, TENS (49); 2. Parity 0, sham TENS (51); 3. Parity 3 TENS (27); 4. Parity 3 sham TENS (23). From admission to labour ward.	Patients assured that they could use TENS until confident; that other forms of analgesia would be available; nature of TENS explained.	Titrate controls as contractions increased, they were not made aware of any sensations at site of electrodes.	3M Tenzcare, dual channel	Individually titrated, pulse width 60–80 µs, repetition rate 80–100/second.	Two pairs: one pair T10–L1 dermatome, 5 cm away from either side of spine; one pair posterior rami S2–L1.	Sham TENS (no current).	Described as double-blind. TENS machine with red light, no current. Neither patient or attending midwife aware which treatments were allocated, third party changed numbers in attempt to maintain blinding.	1
Lee, *et al.*, 1990[13]	Primigravida and second uncomplicated labour; age range 18–35 years.	Parallel group: 1. TENS (38); 2. sham TENS (33); 3. No treatment control (34).	Patients given verbal explanation on rationale for TENS. Additional analgesia given as necessary, but epidurals not encouraged.	Low-frequency TENS was commenced during first stage of labour, high frequency TENS used during contractions and second stage. Patients not made aware of sensations at site of stimulation.	Eastleigh Obstetric device	Biphasic rectangular wave, frequency range 2–200 Hz bursts/second. Amplitude individually titrated by patients.	Two pairs silicone electrodes, 4 x 12 cm; T10–L1 and S2–4 spinal level.	Sham TENS (no current); control (no TENS device).	Neither patients nor obstetric staff knew which devices were active or inactive.	1
Nesheim, 1981[14]	Expected birth following normal labour and normal pregnancy; cervical dilation < 4 cm.	Parallel group: 1. TENS (35); 2. sham TENS (35). Began when dilation < 4 cm.	Aim was to try TENS to determine its effectiveness in labour pain, no risks and alternative methods would be available if PR inadequate.	Two sets of instructions: 1. Explanation of active treatment to patient and partner, encouraged to use when contraction began, to use more intensely as pain increased, free to stop at any time. 2. Sham – to expect no sensation other than warmth.	Travisens, Dan-Sjo Elektronik	Individually titrated and decreased until comfortable, pulse 0–40 mA, frequency 100 Hz, pulse duration 0.25 ms, frequency 50–150 Hz.	Two pairs: both pairs T10–L1.	Sham TENS (no current)	Red light, not blind to observer as patients given different instructions.	2

continued

TABLE 59 contd *Randomised studies of TENS in labour pain*

Study	Early criteria	Study design and duration of treatment periods	Pre-study instructions	Intra-study instructions	TENS machine	TENS setting	Electrode details	Control group(s)	Blinding	Quality score
Steptoe & Bo, 1984[15]	Normal vaginal delivery, primigravida, > 3 cm dilation.	Parallel group: 1. TENS (13); 2. sham TENS (12). Stimulation for 30 minutes.	Identical short verbal and written information to both groups.	Same to both groups to titrate up to level of comfort or adequate PR during first 30 minutes.	Elpha 500	Individually titrated over 30 minutes, 0–60 mA, pulse width 0.2 ms, frequency 1–4 Hz, 100 Hz.	Two pairs carbon rubber: T10–T12, S2–S3.	Sham TENS + standard obstetric analgesia.	Red light, same instruction to both groups.	2
Thomas, *et al.*, 1988[16]	Early labour in primigravida and multigravida, normal or induced delivery, < 7 cm cervical dilation.	Parallel group: 1. TENS (132); 2. sham TENS (148). TENS applied when discomfort reported.	No ante-partum instruction given on TENS. Standard protocol given to both groups by instructor who only advised on TENS use.	Both groups advised to increase TENS settings as needed during contractions; patients free to use other analgesia if required. TENS turned off for two contractions each hour and differences in pain assessed.	3M, dual channel	Individually titrated.	2 pairs electrodes: one pair paravertebrally on either side of spinous process, T10–L1, one pair, S2–4.	Sham TENS (no current).	Good attempt to blind study, both active and sham machines had flashing light, TENS applied by staff not associated with trial, labour managed by non-study staff in normal way. Instructions not given by assessor.	4
van der Ploeg, *et al.*, 1996[17]	First stage of labour, primiparae multiparae, when analgesia requested.	Parallel group: 1. TENS (46); 2. Sham TENS (48). TENS used until cervix fully dilated.	Use of TENS explained to expectant parents by attending physician. Patients supervised until competent in the use of TENS.	Low-frequency TENS used until contractions when high-frequency TENS used (range 1–6). Titrated by partner (mostly). PCA pethidine and promethazine escape analgesia.	Agar GK (Klinerva Holland)	Individually titrated.	2 pairs electrodes, 50 x 100 mm. 1 cm lateral spine, L1–L3 and L4–S1.	Sham TENS; (identical placebo device).	No description of sham TENS other than identical device.	2
Wattrisse, *et al.*, 1993[8]	Primigravida, gestation of at least 38 weeks, < 3 cm cervical dilation, expected normal delivery with extradural analgesia.	Parallel group: 1. epidural (60); 2. epidural + TENS (60). TENS applied with epidural to end of labour.	Not described.	Not described.	Anesthelec MPO2	High frequency, 166 kHz, 1 ms, 20%; low frequency, 83 Hz, 4 ms, 33%.	3 electrodes: 2 posterior mastoid, 1 between eyebrows.	Epidural 0.25% bupivacaine as required, first bolus with fentanyl 100 µg.	None.	2

in which TENS was seen to reduce additional analgesic interventions[11,15] had only 20 and 25 patients in the comparisons compared with 527 patients in the comparisons in the other three trials *(Figure 47)*. In only one study of moderate size[8] had a secondary outcome been judged by us to be positive for cranial TENS. This emphasises how individual small studies may mislead because of the random play of chance.

The choice of outcome measure is an important determinant of how studies are to be judged. If the objective of TENS is to alleviate pain, then it is fair that judgements of its effectiveness are based on prospective subjective measures of pain intensity or relief (primary outcomes), and that these assessments are done at appropriate time points. Retrospective measures of pain are notoriously unreliable. Subsequent need for other analgesic interventions is a secondary outcome measure but one commonly used in these studies. The implications of these results for current practice is that women who are offered TENS are at risk of having their pain inadequately controlled and may experience delays in receiving effective interventions.

This review was restricted to RCTs, unlike that by Reeve and colleagues.[2] RCTs represent the gold standard in clinical trials of efficacy.[19] For nearly 20 years, non-randomised studies have been known to yield larger estimates of treatment effects than studies using random allocation.[20] The size of the overestimation of treatment effect when randomisation is inappropriate can be as much as 40%. In postoperative pain, non-randomised trials of TENS were more likely to show a positive result than randomised trials, with 15 from 17 randomised trials being negative and 17 from 19 non-randomised trials positive.

TABLE 60 *Results of randomised studies of TENS in labour pain*

Study	Pain outcomes (primary outcomes in bold; secondary outcomes in italics)	Results for pain outcomes	Withdrawals and drop-outs	Adverse effects	Significant difference for at least one primary/secondary outcome	Overall judgement
Bunsden, *et al.*, 1982[9]	1. 5-point PI (hourly) 2. Use of any other pain-relieving interventions 3. Duration of labour 4. Questionnaire on day after delivery; abdominal and back pain assessed independently.	Back pain severe: ≤ 5 cm dilated, TENS 3/15, control 5/9; > 5 cm dilated, TENS 1/7, control 5/6. Suprapubic pain severe: ≤ 5 cm dilated, TENS 10/15, control 7/9; > 5 cm dilated, TENS 7/8, control 5/6. Stage 2: pudendal block 13/15 TENS, 7/9 control analgesic.	1 in each group excluded due to subsequent Caesarean section. 1 in each group received epidural because of special problems which were not described.	No specific effect of TENS on foetal heart rate.	No/no	Negative
Champagne, *et al.*, 1984[11]	1. Additional pain relieving interventions.	5/10 required additional analgesic intervention active stimulation, 10/10 control group.	None reported.	Not described.	N/A /yes	Positive secondary
Chia, *et al.*, 1990[10]	1. 3-point PI pre-escape analgesia 2. 3-point PR 3. Time to next analgesic.	NSD. No relief 11% TENS, 50% Entonox, but contractions significantly higher in Entonox group. Additional analgesia not described.	1 woman in Entonox group delivered without further analgesia, and was excluded.	Not described.	No/no	Negative
Harrison, *et al.*, 1986[12]	Research midwife assessments: 1. 5-point PI (hourly) 2. Baseline pain threshold (Mosanto gun) 3. 4-point PR 4. Site of pain.	NSD between TENS and sham TENS for pain or for those requiring extra analgesia (12% TENS, 14% sham TENS). Pain score > 50% at 1 hour: TENS 63/64, sham TENS 55/59. Additional analgesia needed: 57/76 TENS, 58/74 controls.	Not described.	Not described.	No/no	Negative
Lee, *et al.*, 1990[13]	Every 30 minutes: 1. VAS PI 0–10 every 30 minutes 2. Strength of uterine contractions (weak, moderate, strong). Retrospective questionnaire at 24 hours postpartum (1. did patient find TENS helpful or not? (0–3); 2. future use?).	NSD between TENS and sham TENS. Use of additional analgesic interventions (excluding epidurals): 40/62 TENS, 22/35 sham TENS, 28/37 control. NSD between three treatments for 30-minute pain scores.	Not described.	None reported.	No/no	Negative
Nesheim, 1981[14]	1. Overall 5-point PR after childbirth 2. Escape analgesia.	NSD between TENS and sham TENS. Pain free: 1/35 TENS, 0/35 sham TENS. Good relief: 4/35 TENS 5/35 sham TENS. 63 analgesic interventions in TENS, 49 in controls.	Not described.	Not described.	N/A /no	Negative
Steptoe & Bo, 1984[15]	1. VAS PI 0–10 at baseline and 30 minutes after TENS 2. Other analgesic interventions 3. Time of contractions.	No difference in pain measurements. Additional analgesia: 5/12 TENS, 13/13 control group.	1/13 excluded in TENS group due to failed battery in device; 1/13 in control group had Caesarean section but included in analysis.	0/12 TENS; 0/13 sham TENS.	No/yes	Positive secondary
Thomas, *et al.*, 1988[16]	1. VAS PI (hourly) for abdominal and back pain 2. Use of other analgesic interventions 3. Postpartum overall assessment by patient.	NSD between TENS and sham TENS at < 7 cm dilated, 7–10 cm dilated, or during stage 2. No difference in use of other methods of PR (Entonox, pethidine or epidural). Postpartum assessment of excellent/good relief by 29/132 active group.	52/148 control group requested to withdraw compared with 54/132 in TENS group. Only 96 patients could continue with VAS beyond 7 cm and 16 into second stage of labour.	Not described.	No/no	Negative
van der Ploeg, *et al.*, 1996[17]	1. VAS PI (bad = no; good = yes) 2. Other analgesic interventions 3. Patient's impression of medication during labour 4. Would patient choose TENS in future deliveries? 5. Patient's impression of effect of TENS on pain.	NSD between TENS and sham TENS. Mean number of requests for other analgesia: 18.2 TENS, 26.2 sham TENS. Number of times analgesia administered: 5.9 ± 2.32 TENS, 6.5 ± 1.77 sham TENS. Amount of pethidine administered (mg): 60.8 ± 21.6 TENS.	2 refused to take part and received standard analgesia. No other details given.	None reported.	No/no	Negative
Wattrisse, *et al.*, 1993[8]	1. VAS for global pain quality (patient) during labour at 2 hours post delivery 2. Duration of analgesia after first epidural bolus 3. Time between epidural boli.	Quality of analgesia during dilatation and at delivery not different between active and control groups. Duration of first epidural local anaesthetic bolus increased in TENS group by mean of 22 minutes ($p < 0.01$). Time between boli significantly prolonged in TENS.	7 withdrawals, 1 Caesarean section, 2 technical problem with epidural, 2 electrodes fell off, 1 had other treatment. NB: not all described.	Not described.	N/A /yes	Positive secondary

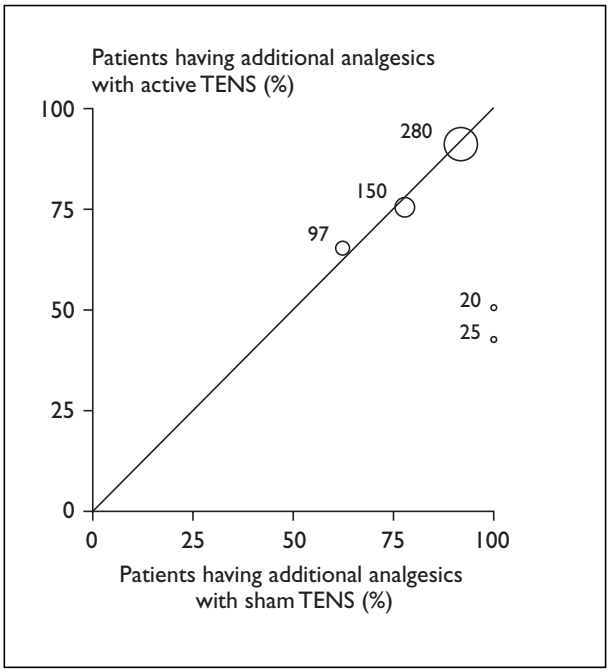


FIGURE 47 *Additional analgesic requirements (numbers of patients in each trial is given next to the circle, whose diameter is proportional to the number)*

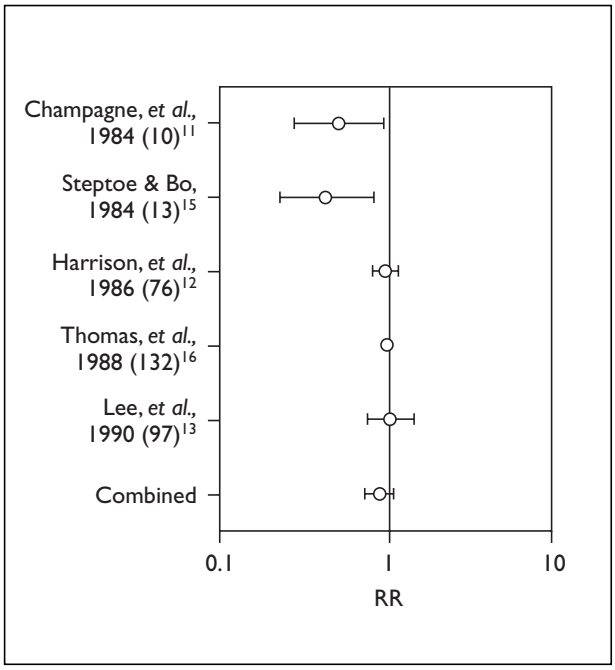


FIGURE 48 *Relative risk for additional analgesic intervention*

The overall methodological quality of the trials reported was low, reflecting the fact that it is difficult, if not impossible, to blind studies of TENS.[21] Inadequate blinding may be an important source of observer bias and may contribute to overestimation of treatment effects. Four of the studies considered here made no attempt at blinding and, of the seven that used sham TENS, only one[16] described the method of blinding in sufficient detail to indicate that blinding may have been adequate.

The 1994 Maternity Service Charter[22] tells women that "...you have the right to be given an explanation of any treatment proposed, including the benefits and risks and of any alternatives before you decide whether you will agree with the treatment". Those involved in the provision of maternity services therefore need to be aware of current research findings concerning effective interventions for the relief of pain, so they can apply these findings in their clinical practice and provide women with accurate information so they can be involved in decisions concerning their care.

On the basis of these findings the continued use of TENS in childbirth needs to be carefully reconsidered. The continued use of TENS in labour pain has considerable implications both for maternity services and the women who use TENS, in terms of receiving prompt and effective pain relief during childbirth. Instead of TENS, women should be given the option of more effective interventions.

References

1. Jadad AR, McQuay HJ. A high-yield strategy to identify randomized controlled trials for systematic reviews. *Online J Curr Clin Trials* [serial online] 1993 Feb 27;Doc No 33:3973 words; 39 paragraphs; 5 tables.
2. Reeve J, Menon D, Corabian P. Transcutaneous electrical nerve stimulation (TENS): a technology assessment. *Int J Technol Assess Health Care* 1996;**12**:299–324.
3. L'Abbé KA, Detsky AS, O'Rourke K. Meta-analysis in clinical research. *Ann Intern Med* 1987;**107**:224–33.
4. Jadad AR, Moore RA, Carroll D, Jenkinson C, Reynolds DJM, Gavaghan DJ, *et al.* Assessing the quality of reports of randomized clinical trials: is blinding necessary? *Control Clin Trials* 1996;**17**:1–12.
5. DerSimonian R, Laird N. Meta-analysis of clinical trials. *Control Clin Trials* 1986;**7**:177–88.
6. Cook RJ, Sackett DL. The number needed to treat: a clinically useful measure of treatment effect. *BMJ* 1995;**310**:452–4.
7. Carroll D, Tramèr M, McQuay H, Nye B, Moore A. Transcutaneous electrical nerve stimulation in labour pain: a systematic review. *Br J Obstet Gynaecol* 1997;**104**:169–75.
8. Wattrisse G, Leroy B, Dufossez F, Bui Huu Tai R. Electrostimulation cerebrale transcutanee: étude comparative des effets de son association a l'anesthesie peridurale par bupivacaine-fentanyl au cours de l'analgesie obstetricale. *Cah Anesthesiol* 1993;**41**:489–95.

9. Bundsen P, Ericson K, Peterson LE, Thiringer K. Pain relief in labor by transcutaneous electrical nerve stimulation. Testing of a modified stimulation technique and evaluation of the neurological and biochemical condition of the newborn infant. *Acta Obstet Gynecol Scand* 1982;**61**:129–36.

10. Chia YT, Arulkumaran S, Chua S, Ratnam SS. Effectiveness of transcutaneous electric nerve stimulator for pain relief in labour. *Asia Oceania J Obstet Gynaecol* 1990;**16**:145–51.

11. Champagne C, Papiernik E, Thierry JP, Noviant Y. Electrostimulation cerebrale transcutanee par les courants de Limoge au cours de l'accouchement. *Ann Fr Anesth Reanim* 1984;**3**:405–13.

12. Harrison RF, Woods T, Shore M, Mathews G, Unwin A. Pain relief in labour using transcutaneous electrical nerve stimulation (TENS). A TENS/TENS placebo controlled study in two parity groups. *Br J Obstet Gynaecol* 1986;**93**:739–46.

13. Lee EWC, Chung IWY, Lee JYL, Lam PWY, Chin RKH. The role of transcutaneous electrical nerve stimulation in management of labour in obstetric patients. *Asia Oceania J Obstet Gynaecol* 1990;**16**:247–54.

14. Nesheim BI. The use of transcutaneous nerve stimulation for pain relief during labor. A controlled clinical study. *Acta Obstet Gynecol Scand* 1981;**60**:13–16.

15. Steptoe P, Bo JO. Transkutan nervestimulations smertelindrende effekt ved fødsler. *Ugeskr Laeger* 1984;**146**:3186–9.

16. Thomas IL, Tyle V, Webster J, Neilson A. An evaluation of transcutaneous electrical nerve stimulation for pain relief in labour. *Aust N Z J Obstet Gynaecol* 1988;**28**:182–9.

17. van der Ploeg JM, Vervest HAM, Liem AL, van Leeuwen S. Transcutaneous nerve stimulation (TENS) during the first stage of labour: a randomized clinical trial. *Pain* 1996;**68**:75–8.

18. Carroll D, Moore RA, Tramèr MR, McQuay HJ. Transcutaneous electrical nerve stimulation does not relieve labour pain – updated systematic review. *Curr Rev Obstet Gynaecol* 1997;(Sept):195–205.

19. Schulz KF, Chalmers I, Hayes RJ, Altman DG. Empirical evidence of bias: dimensions of methodological quality associated with estimates of treatment effects in controlled trials. *JAMA* 1995;**273**:408–12.

20. Chalmers TC, Matta RJ, Smith J Jr, Kunzler AM. Evidence favoring the use of anticoagulants in the hospital phase of acute myocardial infarction. *N Engl J Med* 1977;**297**:1091–6.

21. Deyo RA, Walsh NE, Schoenfeld LS, Ramamurthy S. Can trials of physical treatments be blinded? The example of transcutaneous electrical nerve stimulation for chronic pain. *Am J Phys Med Rehabil* 1990;**69**:6–10.

22. The patient's charter: maternity services. London: Department of Health, 1994.

Chapter 21

Efficacy and harm of anti-emetic interventions in the surgical setting

The project described in this chapter had the aim of establishing a league table of the relative prophylactic efficacy and the likelihood of harm from anti-emetic interventions which are currently used to treat or prevent PONV.

Data source

Systematic reviews of the literature, using data extraction from relevant reports, critical appraisal of data and meta-analytical combination of data with biostatistical methods, were chosen as the main instrument for this study. Thus the 'essence' of this work comes from ten different systematic reviews, of which eight were designed to investigate the efficacy and potential for harm of anti-emetic interventions. At least 860 reports were screened, from which 215 RCTs with data from 31,801 patients were analysed. The two remaining systematic reviews were designed to investigate harm from interventions. For this purpose, after screening about 1200 reports, 300 reports of different study architectures with data from 1,432,817 patients were analysed.

Efficacy of anti-emetic interventions

A model was proposed to compare anti-emetic interventions indirectly, that is, without the need for direct comparisons,[1] which is formally equivalent to that used to generate the league table for analgesic interventions. There is one difference, however. When placebo responses for analgesics vary, they tend to vary around a population placebo event rate of about 19%, that is, about 19% of patients with moderate or severe acute pain given a placebo will experience at least 50% pain relief. For PONV, the spread is much greater, with percentages of patients vomiting without prophylactic interventions in studies ranging from close to zero to over 80%, and there seems to be no central figure or population response. This is the case even when operation, anaesthetic and patient population are highly standardised, as in the case of paediatric strabismus surgery. Clearly PONV is different from postoperative pain in this respect.

There is a validity issue which has been addressed by limiting the range of included trials to those which duplicate usual clinical experience. Two narrow bands of CER were therefore defined: 20–60% CER for early outcomes (0–6 hours after surgery) and 40–80% CER for late outcomes (0–48 hours). Only trials with CERs within these bands were analysed. Validity criteria defined lower and upper boundaries. Trials with very low CERs do not allow a valid assay of anti-emetic efficacy; PONV cannot be prevented if nobody is going to vomit without prophylaxis. Trials with very high CERs do not represent daily clinical practice but enable even marginally active interventions to show statistically significant efficacy. Trials with very low or very high CERs were, therefore, regarded as invalid. The arbitrary limits of the CER banding were applied to all analysed interventions.

The league table of anti-emetic efficacy

Five prophylactic interventions were analysed within the CER banding, each with data from systematically searched RCTs: propofol induction (reworking of data from Tramèr and colleagues[2]), propofol maintenance,[1] omitting nitrous oxide,[3] a total intravenous anaesthetic with propofol,[1] and ondansetron.[4] Propofol was compared with non-propofol anaesthetics, which were regarded as 'no treatment' controls. Omitting nitrous oxide was compared with using it ('no treatment' control). In ondansetron trials, comparators were placebos or no treatment. Results are presented in *Figures 49–52.*

Interpreting the league table of efficacy

Four criteria need to be taken into account.

Endpoint

The endpoint is prevention of nausea or vomiting, within 6 hours of surgery (early efficacy) or within 48 hours of surgery (late efficacy). Thus for each intervention, anti-nausea and anti-vomiting efficacy can be interpreted separately, as well as short-term and long-term efficacy.

Short-term efficacy has an economic impact mainly in day-case surgery where patients are meant to be discharged within hours of the procedure; they

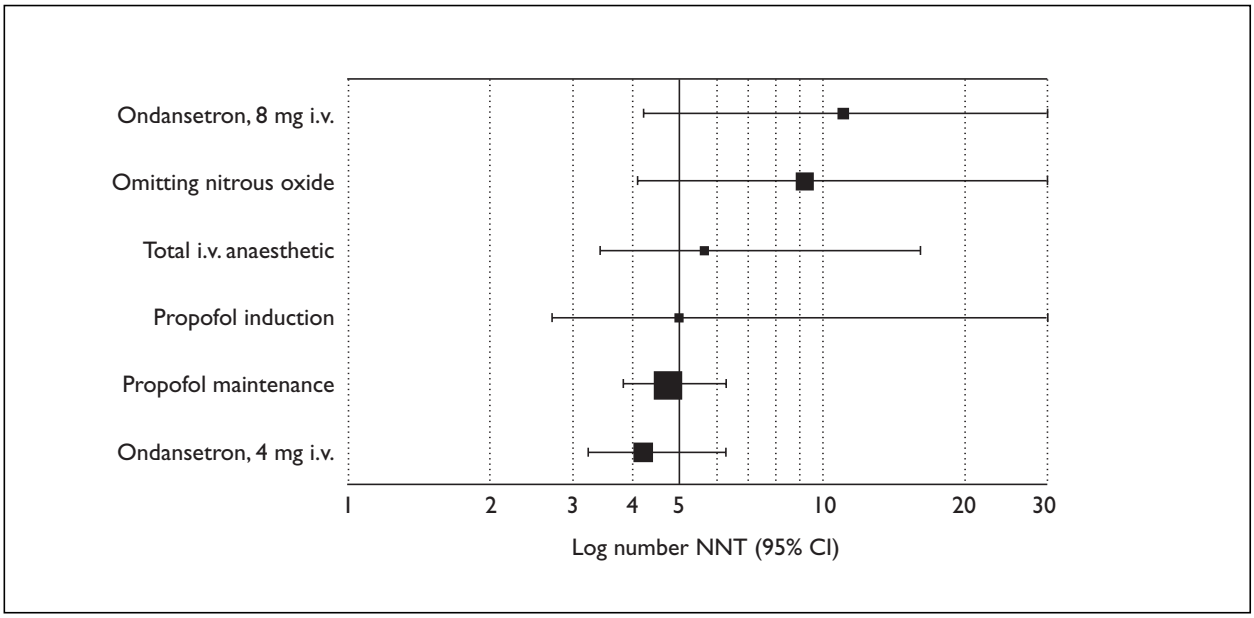


FIGURE 49 *Prevention of early nausea (0–6 hours): CER 20–60%*

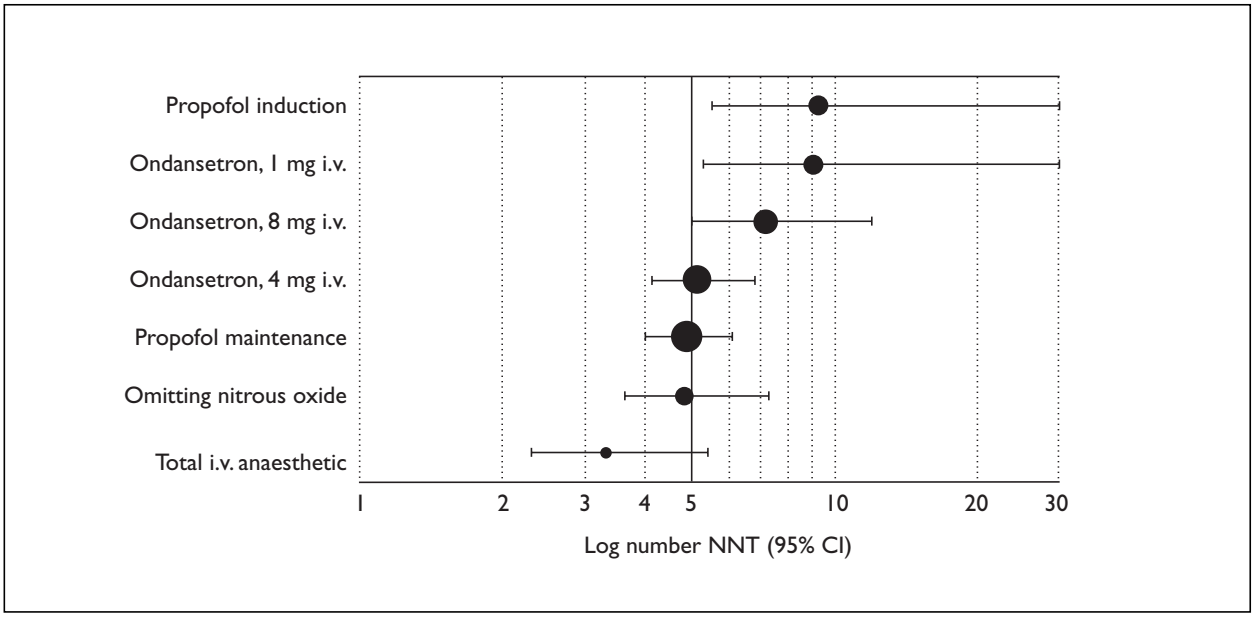


FIGURE 50 *Prevention of early vomiting (0–6 hours): CER 20–60%*

have to be free of PONV in order to fulfil discharge criteria. Long-term efficacy is a better indicator of the drug's anti-emetic efficacy and patients' comfort. It indicates if the patient will remain PONV-free at home (or on the journey home).

NNT point estimate

Worthwhile anti-emetic efficacy in the surgical setting was arbitrarily defined as an NNT to prevent nausea or vomiting compared with placebo (or no treatment) of ≤ 5.[2] This means that at least 20% of treated patients will profit from the prophylaxis. The lower boundary of the CER banding for early outcomes was set at 20%.[1] Thus, interventions which had no scope to show an NNT of at least 5 for efficacy were not considered in the model.

Confidence interval

The upper boundary of the CI for the NNT places the treatment in the least favourable light. If this upper limit lies within what would be considered to be the minimal clinically relevant

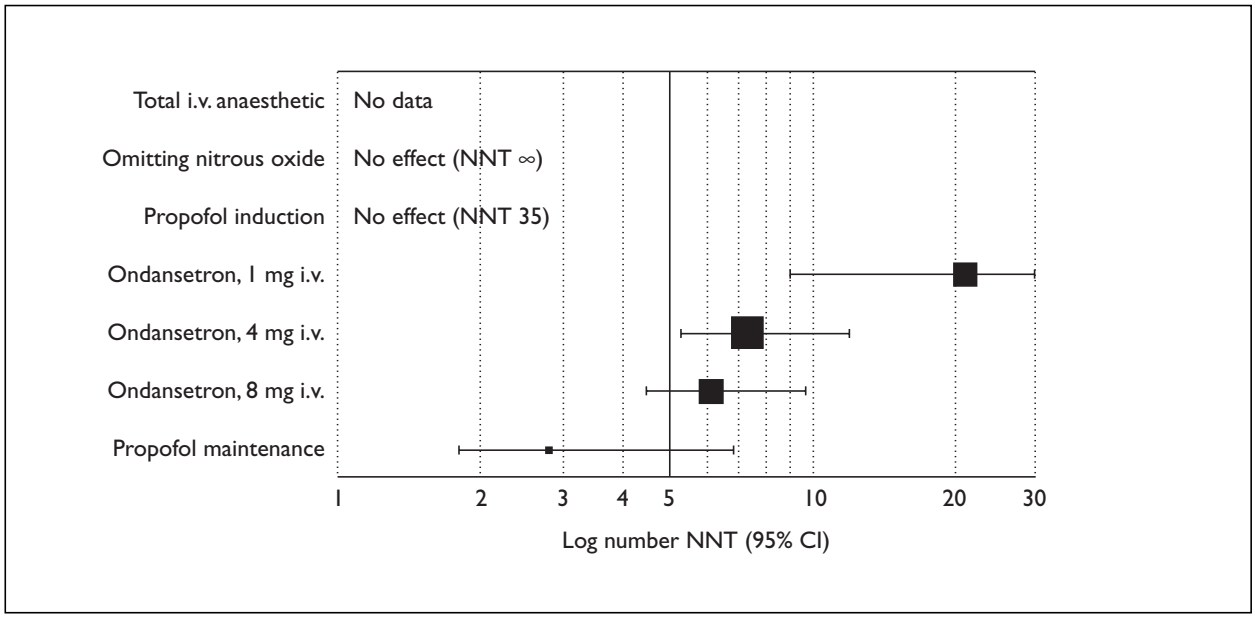


FIGURE 51 *Prevention of early nausea (0–6 hours): CER 20–60%*

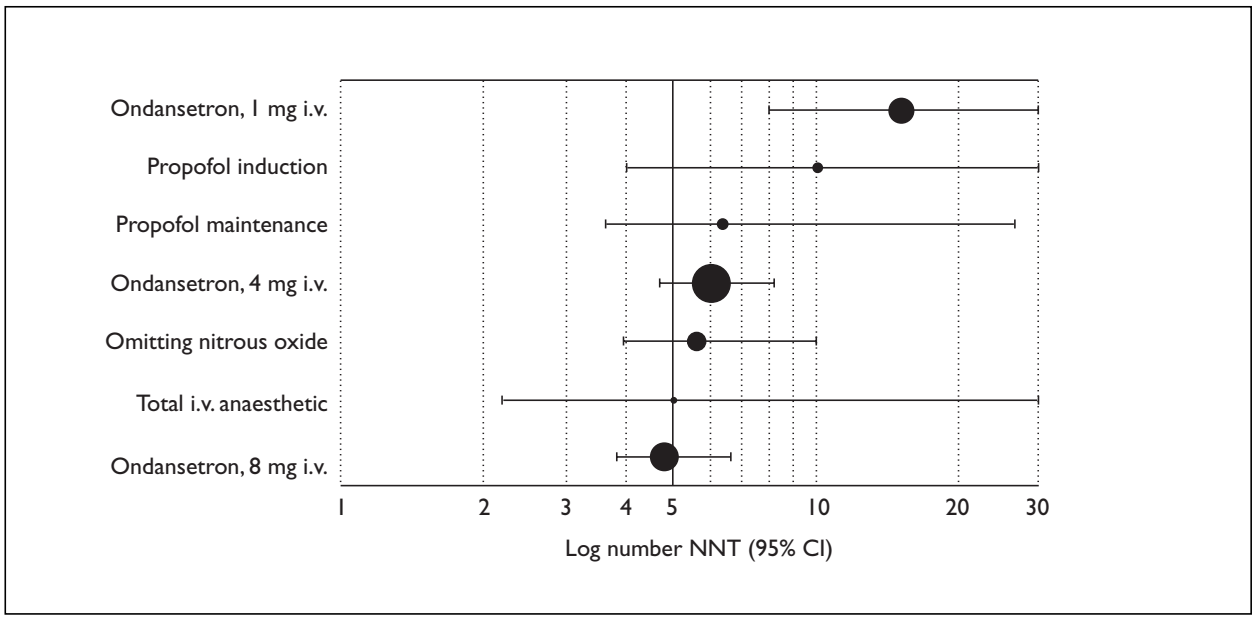


FIGURE 52 *Prevention of early vomiting (0–6 hours): CER 20–60%*

effect (for instance, an NNT of 5 to prevent PONV),[2] the result indicates a definitely useful treatment. Lack of overlap between CIs can be used as a simple test of the statistical difference between NNTs of two interventions.

Size (area) of the symbol

Areas of symbols were plotted that were proportional to the number of analysed patients. The larger the number of analysed patients (that is, the larger the symbol), the greater the confidence in the point estimate.

Results

The best anti-emetic prophylaxis has the lowest NNT (but not above 5), the largest symbol, the narrowest CI (the upper limit below 5), and the most consistent efficacy (that is, anti- both nausea and vomiting, and both short- and long-term). No anti-emetic intervention which has been tested meets all these criteria. Intravenous ondansetron, 8 mg, comes closest. Propofol maintenance looks promising but the long-term effect is based on a limited number of patients only; the symbol is small and the

CI large. Omitting nitrous oxide looks promising too but only for anti-vomiting efficacy.

An intervention which indicates good efficacy (i.e. NNT < 5) but which is based on a limited number of patients and, therefore, with a small symbol and/or a wide CI, may lead to a research agenda. The question then is: 'Are further trials needed?' and, if the answer is affirmative, 'What trials are needed?'

Potential for harm of anti-emetic interventions

Interpretation of any intervention's clinical usefulness must take into account both efficacy and harm. An objective measurement of severity does not exist; acceptability and interpretation of harm are likely to be multifactorial. Unlike the efficacy league table, the league table on intervention-related harm is not based on truncated data sets. All comparisons are between the active interventions and placebo or no treatment or in, the case of propofol, between active (propofol) and another, propofol-free anaesthetic (control).

The league table of harm

The league table of harm presented in *Figure 53* is based on analyses of four systematic reviews: droperidol,[5] omitting nitrous oxide,[3] propofol,[6] and ondansetron.[4]

Interpreting the league table of harm

Endpoint

A specific endpoint and its risk have to be interpreted in their proper context. For instance, elevated liver enzymes with ondansetron might be perceived as trivial. The biological basis for this adverse drug reaction, however, is not known. There may be an argument for not giving ondansetron to patients with pre-existing abnormal liver function tests or underlying liver disease. It is unknown if ondansetron should be avoided when other potentially hepatotoxic drugs, such as halothane or paracetamol, are to be used.

NNH point estimate

The NNH indicates in how many patients the adverse drug reaction will occur which would not have happened had the patient not received the drug. Interpretation of this result has to take into account other factors, such as severity of the reaction, severity of the underlying disease, and availability of alternative treatments and their potential for harm.

Statistical significance

In contrast to the league table of anti-emetic efficacy, symbol sizes are not plotted proportional to the quantity of analysed data (i.e. the symbol area is fixed) and no CIs are shown. The reason is that some NNHs are based on a limited number of patients who had the adverse drug reaction. Yet these NNHs may be clinically relevant. Plotting

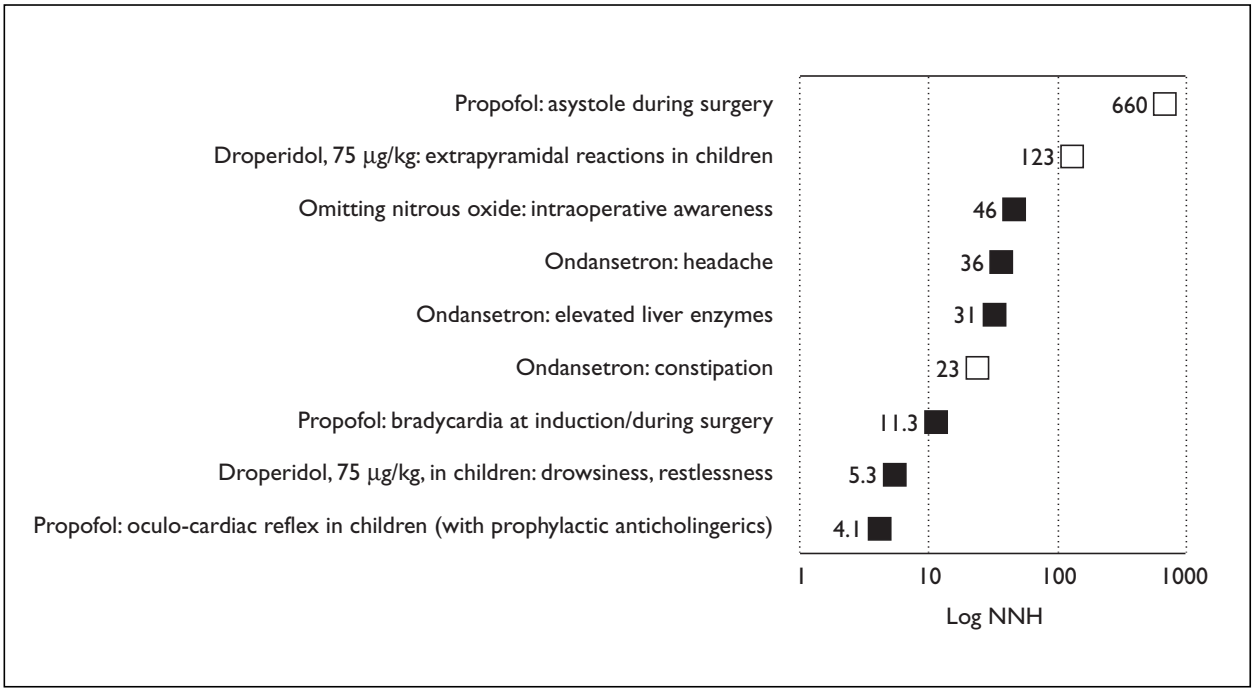


FIGURE 53 *Harm with anti-emetic interventions. All data are from RCTs. Significance = statistical significance, assumed when* p < *0.05*

symbol size in relation to the quantity of analysed data would detract from such potentially important results and tend to overinterpret more trivial (but better documented) reactions.

A graphical distinction was made between significant and non-significant findings. Statistical significance was arbitrarily set at a value of $p < 0.05$. This value may be unnecessarily conservative and narrower CIs could be chosen. Black symbols represent adverse drug reactions which happened statistically significantly ($p < 0.05$) more often with the intervention. Accordingly, white symbols indicate absence of statistical significance ($p > 0.05$).

Conclusions

Each intervention introduces a certain risk of adverse drug reactions. The league table of intervention-related harm is an important contributor to a rational risk–benefit assessment. It will help both physicians and patients to take decisions about the use (prophylactically or therapeutically) of anti-emetic interventions.

Prophylaxis versus treatment

One of the main conclusions of the first systematic review of efficacy and harm of anti-emetic interventions in the surgical setting[5] was that it would perhaps be better to wait and see which patient vomits and then treat. This conclusion was based on the somewhat unexpected result that even in paediatric strabismus surgery, a clinical setting with a high risk for PONV, only 25% of the children actually profited from the best prophylaxis. The subsequently calculated NNTs of all the other prophylactic interventions (propofol induction and maintenance, omitting nitrous oxide, ondansetron) did not prove to be more efficacious.

The justification of prophylactic postoperative antiemetics was queried 35 years ago by Adriani and colleagues.[7] They noted that no more than 25% of patients in the recovery room vomited in the immediate postanaesthesia period, that most of this vomiting was short-lived and subsided spontaneously without the use of antiemetics. Similar average PONV incidence has subsequently been reported repeatedly in large RCTs, in case series and in systematic reviews of RCTs, although it may be higher in specific clinical settings, such as paediatric strabismus surgery. If the incidence without prophylaxis is only 25% and treatment is effective, then arguably prophylaxis may be unnecessary on grounds of adverse effects and cost. The humanitarian argument is that it is unacceptable to wait and see if a patient is going to vomit or become nauseated before starting treatment. Also there is a widespread belief that it may be more difficult to treat established PONV than to prevent it,[8] although there is no substantial evidence to support this view. Using a decision-analysis treatment model, it has been suggested that prophylactic anti-emetic therapy was more cost-effective compared with treatment of established symptoms for operations associated with a high incidence of emesis.[9]

The pivotal evidence to resolve the debate was the relative effectiveness of treatment and prophylaxis of PONV. This comparison was possible with ondansetron. The finding of a dose–response relationship with ondansetron in prevention of PONV[4] contrasted with the analysis of the efficacy of ondansetron in the treatment of established PONV.[10] For this, no dose response between 1 mg and 8 mg could be established; the NNT to prevent further PONV in a nauseated or vomiting patient with the lowest dose tested, 1 mg, compared with placebo was between 4 and 5, and higher doses were no more effective. Thus, 1 mg is as efficacious for treatment of established PONV as an eight-fold higher dose (i.e. 8 mg) is for the prevention of PONV.[4] This challenges the usefulness of prophylactic ondansetron when risk–benefit and cost–benefit arguments are considered.

Two scenarios

Two scenarios may be described that illustrate the relationship between prophylaxis and treatment based on these numbers. For simplicity, several assumptions have to be made:

(i) one ondansetron, 2 mg, ampoule is set at £6.75 and a 4 mg ampoule at £13.50[11]
(ii) one ondansetron, 2 mg, ampoule (the smallest commercially available) per patient will be used to treat established PONV; two 4 mg ampoules per patient will be used for prophylaxis
(iii) ondansetron's prophylactic anti-emetic efficacy is independent of CER, although its prophylactic efficacy is likely to be lower (i.e. higher NNT) with low CERs
(iv) the success rate with ondansetron in patients with established PONV is 40%,[10] which means that 40% of all vomiting or nauseated patients receiving ondansetron will not vomit or be nauseated further
(v) NNTs for the optimal ondansetron doses are set at 5 for both prevention and treatment (instead of 5–6 for prevention and 4–5 for treatment)

(vi) NNH is set at 30 for both prevention and treatment, and there is no dose response (i.e. the risk of an adverse drug reaction is similar for both 2 mg and 8 mg); the adverse drug reaction could be a headache or elevated liver enzymes.

Any calculation based on these assumptions will yield an overoptimistic result for prevention and a conservative result for treatment of PONV with ondansetron.

In the first scenario, for 1000 'high risk' patients undergoing surgery the CER is 60% (see *Box 3).*

In this situation, prophylaxis costs (1000 x (2 x £13.50)) are £27,000, whereas treatment costs (600 x £6.75) are £4050. The benefit is £22,950, or £22.95 per patient. In the prophylaxis group, 400 patients are failures (i.e. they are nauseated or they vomit despite prophylaxis) compared with 360 in the treatment group (i.e. they continue to be nauseated or to vomit because the treatment failed). Almost twice as many patients receiving the drug prophylactically will have an adverse drug reaction.

In the second scenario, a 'low risk' situation, 1000 patients undergo surgery and the CER is 30% (see *Box 3*).

Cost is maintained at £27,000 in the prophylaxis group but decreases to £2025 (300 x £6.75) in the treatment group, a benefit of £24,925, or almost £25 per patient. In the prophylaxis group, 100 patients are failures, compared with 180 in the treatment group who continue to be nauseated or to vomit. More than three times as many patients in the prophylaxis group have an adverse drug reaction compared with patients who are treated for established PONV.

These risk–benefit–cost calculations are the strongest argument against prophylactic use, and in favour of therapeutic use of anti-emetic interventions in the surgical setting. The main arguments are that costs with treatment are dramatically lower and about half as many patients will have an adverse drug reaction. The number of failures is comparable in the high-risk scenario but in favour of prophylaxis in the low-risk scenario. The efficacy of ondansetron or other anti-emetics in patients who already have received ondansetron and in whom the prophylaxis or treatment failed is unknown. The 'price' patients in the treatment group have to pay, is that they need to vomit or feel nauseated before they receive treatment.

The cost-effectiveness arguments for prophylaxis versus treatment are investigated more fully in the next chapter.

BOX 3 Risk–benefit–cost analysis of anti-emetic prophylaxis versus treatment in a clinical setting with high and low CERs

High CER

1000 patients undergoing surgery: CER = 60%

Prevention in 1000 patients: ondansetron, 8 mg i.v. (NNT ± 5)		Treatment of 600 patients with one episode of PONV: ondansetron, 1 mg i.v. (NNT ± 5)
1000 x 8000 mg	**x 13**	600 x 1 mg = 600 mg
400 no PONV anyway + 200 successful preventions		400 no PONV anyway + 240 successful treatments
400 failures		360 failures
Adverse drug reaction in 33 (NNH 30)		**Adverse drug reaction in 20 (NNH 30)**

Low CER

1000 patients undergoing surgery: CER = 30%

Prevention in 1000 patients: ondansetron, 8 mg i.v. (NNT ± 5)		Treatment of 300 patients with one episode of PONV: ondansetron, 1 mg i.v. (NNT ± 5)
1000 x 8 mg = 8000 mg	**x 27**	300 x 1 mg = 300 mg
700 no PONV anyway + 200 successful preventions		700 no PONV anyway + 120 successful treatments
100 failures		180 failures
Adverse drug reaction in 33 (NNH 30)		**Adverse drug reaction in 10 (NNH 30)**

Research agenda

Further systematic reviews on the efficacy and harm of the other anti-emetic drugs which are currently used in anaesthesia and surgery are needed. Drugs of interest include droperidol, metoclopramide and hyoscine (scopoderm). New 5-HT_3 receptor antagonists are increasingly used (granisetron, tropisetron, dolasetron). Dose–response relationships, optimal doses and adverse effect profiles of these drugs have to be established. Opioid-induced nausea and vomiting, for instance, related to PCA or spinal opioid administration, remains a particular problem. The proposed banding model may be used for indirect comparisons of relative anti-emetic efficacy. The risk–benefit–cost model could be used for further process analysis as new data on prevention and treatment with other drugs becomes available. Systematic reviews are likely to inform our future clinical management. However, such projects are time-consuming, expensive, and dependent on efficacious team-work. Interest groups will need to be formed.

References

1. Tramèr M, Moore A, McQuay H. Meta-analytic comparison of prophylactic antiemetic efficacy for postoperative nausea and vomiting: propofol anaesthesia vs omitting nitrous oxide vs a total i.v. anaesthesia with propofol. *Br J Anaesth* 1997;**78**:256–9.
2. Tramèr M, Moore A, McQuay H. Propofol anaesthesia and postoperative nausea and vomiting: quantitative systematic review of randomized controlled studies. *Br J Anaesth* 1997;**78**:247–55.
3. Tramèr M, Moore A, McQuay H. Omitting nitrous oxide in general anaesthesia: meta-analysis of intraoperative awareness and postoperative emesis in randomized controlled trials. *Br J Anaesth* 1996;**76**:186–93.
4. Tramèr MR, Reynolds DJM, Moore RA, McQuay HJ. Efficacy, dose-response, and safety of ondansetron in prevention of postoperative nausea and vomiting: a quantitative systematic review of randomised placebo-controlled trials. *Anesthesiology* 1997;**87**:1277–89.
5. Tramèr M, Moore A, McQuay H. Prevention of vomiting after paediatric strabismus surgery: a systematic review using the numbers-needed-to-treat method. *Br J Anaesth* 1995;**75**:556–61.
6. Tramèr MR, Moore RA, McQuay HJ. Propofol and bradycardia: causation, frequency and severity. *Br J Anaesth* 1997;**78**:642–51.
7. Adriani J, Summers FW, Antony SO. Is the prophylactic use of antiemetics in surgical patients justified? *JAMA* 1961;**175**:666–71.
8. Zahl K. The role of ondansetron and other antiemetics in ambulatory surgery. *J Clin Anesth* 1993;**5** suppl 1:52–56S.
9. Watcha M, Smith I. Cost-effectiveness analysis of antiemetic therapy for ambulatory surgery. *J Clin Anesth* 1994;**6**:370–7.
10. Tramèr MR, Moore RA, Reynolds DJM, McQuay HJ. A quantitative systematic review of ondansetron in treatment of established postoperative nausea and vomiting. *BMJ* 1997;**314**:1088–92.
11. Drugs used in nausea and vertigo (chapter 4.6). British National Formulary, no. 32, 1996. London: British Medical Association and Royal Pharmaceutical Society of Great Britain, 1996.

Chapter 22

Cost-effectiveness of ondansetron: prophylaxis compared with treatment in PONV

Summary

PONV is a frequent complication of anaesthesia and surgery. There is no general agreement as to whether prophylaxis of PONV is better than treatment of established PONV. This lack of agreement may be due to the great variation in the proportion of patients who experience PONV. Given this equipoise, choice of prophylaxis or treatment could be determined by substantial cost-effectiveness difference between the two approaches.

Cost-effectiveness was calculated for ondansetron as prophylaxis or as treatment of PONV using information from systematic reviews and published meta-analyses. Modelling, which was based on a cohort of 1000 patients, examined CERs (vomiting in the absence of prophylaxis) between 10% and 90%, and different doses. Anti-emetic efficacy was assumed to be constant across CERs. In a sensitivity analysis, cost-effectiveness of recommended doses (4 mg for both treatment and prophylaxis) was compared with cost-effectiveness of the most effective doses as demonstrated by meta-analysis (1 mg for treatment and 8 mg for prophylaxis).

For all CERs fewer patients will suffer any PONV symptom (nausea and or vomiting/retching) at any time after surgery with prophylaxis compared with treatment. However, with both effective treatment doses, 1 mg and 4 mg, fewer milligrams are required for each patient who suffers at most one episode of PONV, compared with 4 mg or 8 mg for prophylaxis. For the endpoint of maintaining a patient PONV-free throughout, treatment with 1 mg was still the most cost-effective, followed by treatment with 4 mg but only at a CER below 80%.

Fewer patients will experience any PONV symptoms with prophylaxis compared with treatment. But prophylaxis was not much more effective than treatment, and treatment of established PONV with effective doses (i.e. 1 mg or 4 mg) is more cost-effective than prophylaxis with effective doses (i.e. 4 mg or 8 mg). This is because of the high success rate with the lowest dose tested (1 mg) in established PONV and the disappointing anti-nausea effect of prophylactic ondansetron even at an eight-fold higher dose.

Introduction

The aim was to assess the relative cost-effectiveness of strategies for treating PONV with ondansetron. The intention was to establish cost-effectiveness relationships based on the strongest evidence currently available. Data from two quantitative systematic reviews of published valid RCTs of ondansetron in the surgical setting were used.[1,2]

Methods

Decision tree, endpoints, and estimates of efficacy

The two strategies to deal with PONV, treatment versus prophylaxis, were displayed graphically as a decision tree *(Figure 54)*.

Treatment arm

Some patients will have no PONV symptoms at any time (T1) (see *Figure 54)*. Success with treatment (T2) was defined as a nauseated or vomiting patient who had no further episode of nausea or vomiting after one dose of ondansetron. These patients vomited or felt nauseated once before they received ondansetron. A treatment failure (T3) was a vomiting or nauseated patient who continued to vomit or to feel nauseated despite treatment with ondansetron.

Anti-emetic efficacy of 1 mg of ondansetron for treatment of established PONV was shown not to be different from 4 mg or 8 mg.[1] However, the optimal dose recommended by the manufacturer was 4 mg. For the purpose of this study, therefore, both 1 mg and 4 mg doses were chosen for sensitivity cost-effectiveness analyses of treatment. Success rate was set *(Table 61)* at 40% for the 1 mg dose and at 45% for the 4 mg dose.[1] This meant that 40% and 45% of nauseated or

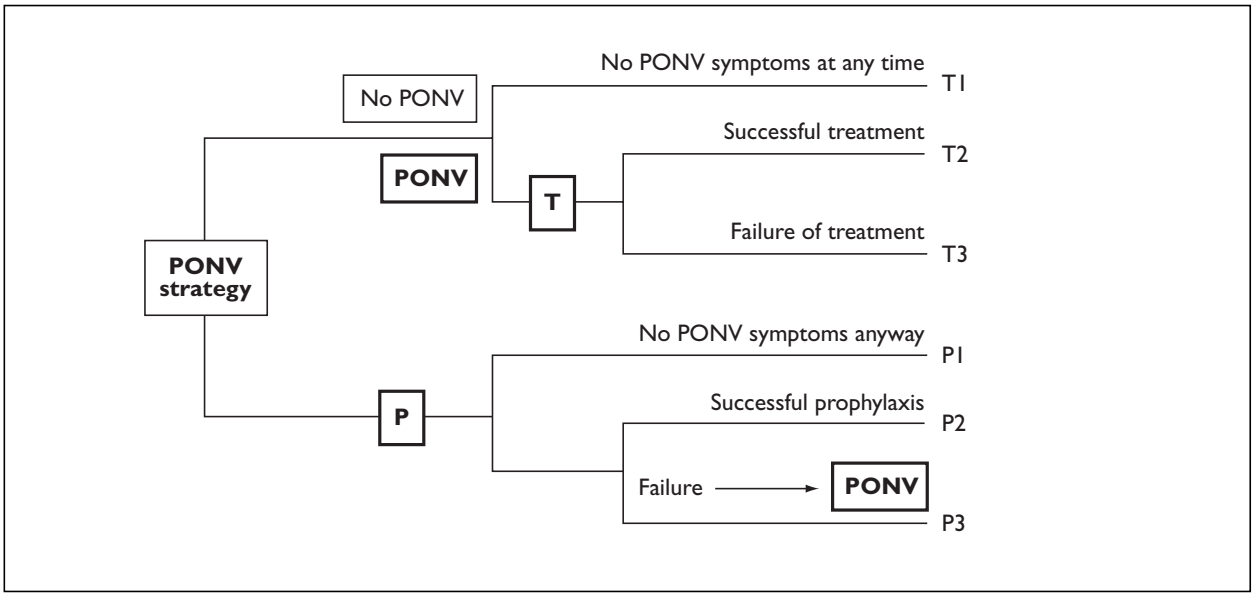


FIGURE 54 *Strategies for dealing with PONV (T, treatment; P, prophylaxis)*

TABLE 61 *Ondansetron: total estimates of efficacy and harm*

Definition	Value	Reference
Success rate with 1 mg treatment	40%	Tramèr, *et al.*, 1997[1]
Success rate with 4 mg treatment	45%	Tramèr, *et al.*, 1997[1]
Success rate with 8 mg treatment	44%	Tramèr, *et al.*, 1997[1]
NNT of 1 mg to prevent nausea	21	Tramèr, *et al.*, 1997[2]
NNT of 1 mg to prevent vomiting	15	Tramèr, *et al.*, 1997[2]
NNT of 4 mg to prevent nausea	16	Tramèr, *et al.*, 1997[2]
NNT of 4 mg to prevent vomiting	6.4	Tramèr, *et al.*, 1997[2]
NNT of 8 mg to prevent nausea	6.4	Tramèr, *et al.*, 1997[2]
NNT of 8 mg to prevent vomiting	5	Tramèr, *et al.*, 1997[2]
NNH for headache with any dose	36	Tramèr, *et al.*, 1997[2]
NNH for elevated liver enzymes with any dose	31	Tramèr, *et al.*, 1997[2]

Success rate = percentage of vomiting or nauseated patients who are treated with the respective dose of ondansetron and who do not continue vomiting or being nauseated.

Success rate with placebo in these trials was 20%.

vomiting patients will not experience any further PONV symptoms (nausea or retching or vomiting) after administration of ondansetron, 1 mg and 4 mg, respectively. It was assumed that the success rate with either dose was independent of both clinical setting and patient. The placebo response of about 20% found in the original systematic review[1] was not taken into account in the analysis.

Prophylaxis arm

Some patients who receive prophylactic antiemetics would not have vomited anyway (P1) (see *Figure 54*). Success with PONV prophylaxis (P2) was represented by a patient who never experienced any PONV symptoms because he or she had received prophylactic ondansetron *(Figure 54)*. A failure in the prophylaxis arm was represented by a patient who received prophylactic ondansetron but

nevertheless vomited or felt nauseated after surgery (P3). For prophylaxis, the appropriate estimate of efficacy was the NNT. This indicated how many patients would have to be treated prophylactically with ondansetron in order to prevent PONV in one patient who would have vomited or been nauseated had placebo been given. Thus the placebo response was taken into account.

The optimal prophylactic dose as recommended by the manufacturer is 4 mg. Meta-analysis, however, showed that 8 mg was the most effective prophylactic dose.[2] Sensitivity analyses to compare cost-effectiveness were therefore undertaken with both the 4 mg and 8 mg doses. The effect of ondansetron on vomiting was consistently more pronounced than its effect on nausea.[2] However, as prophylaxis of nausea may be regarded as being as important as prophylaxis of vomiting, the NNT for prevention of nausea was chosen as the appropriate estimate of efficacy for prophylactic doses *(Table 53):* 16 for a 4 mg dose and 6.4 for 8 mg.[2] It was assumed that ondansetron's prophylactic anti-emetic efficacy was independent of the CER.

Estimate of harm

The NNH was regarded as the appropriate estimate of the likelihood for drug-related adverse effects *(Table 53)*. The NNH was assumed to be 30 for both prophylaxis and treatment with ondansetron.[2] The assumption was made that there was no dose– response (i.e. the additional risk for an adverse drug reaction was similar for 1 mg treatment and 8 mg prophylaxis). The adverse drug reaction could be a headache or elevated liver enzymes.

Modelling

The total cost of arriving at each of the endpoints (P1–P3; T1–T3) *(Figure 54)* was calculated for both strategies, treatment and prophylaxis, based on a cohort of 1000 patients. This was calculated for increasing CERs, from 10% to 90%, and different doses of ondansetron (1 mg and 4 mg for treatment, 4 mg and 8 mg for prophylaxis). Cost-effectiveness ratios were calculated and displayed graphically.

Relevant ratios were:

(i) cost per patient
(ii) cost per patient who experienced no more than one episode of PONV
(iii) cost per patient who never suffered any PONV symptoms.

Such ratios represent the average cost-effectiveness ratios, which are simply the costs of generating the desired endpoint (a PONV-free patient, for instance) divided by the number of patients involved.

Incremental cost-effectiveness ratios were calculated to indicate how much it costs to produce the additional effects.[3] Starting from the least effective strategy, the difference between the costs of pairs of strategies was divided by the difference in their effectiveness.

Drug acquisition costs differ widely because hospitals have different purchasing strategies. The main cost parameter was the number of milligrams of ondansetron required rather than the actual price. Thus, milligram outcomes as reported in this paper may be multiplied by the price per milligram to generate the actual drug costs of each strategy. BNF costs (1997) for ondansetron were £6.75, £13.50 and £27.00 for 2, 4 and 8 mg ampoules, giving an approximate price of £3.40 per milligram.

For the same reason, wide differences between and within hospitals, expenses involved in administration of the drug, cleaning, extra staff time and materials used, costs for rescue anti-emetic medication, and costs for unscheduled admission due to prolonged PONV were not considered. The sum of these supplementary costs may be regarded as a hospital-specific constant which may then be added to the reported costs. Costs which may be relevant would be staff time for administering a treatment to a patient with PONV (say 3 minutes per patient at £15 per hour) and materials (mainly drugs, so say £1 on the basis that older and cheaper drugs than those in the ondansetron class would be used). Any major costs would derive from re-admission or non-discharge costs assumed to be that of an in-patient day cost (£200).

Endpoints

Two endpoints were considered to be particularly important: the number of patients who suffered no more than a single episode of PONV, and the number of patients who suffered no PONV symptoms.

In the treatment arm, the number of patients who suffered no more than one PONV episode were patients who did not vomit or feel nauseated at all (i.e. who did not need any treatment) or, when having symptoms, responded promptly to treatment *(Figure 54:* T1 + T2). This more pragmatic approach would thus accept that a patient vomits once or feels nauseated briefly before an effective treatment is administered. The second, stricter endpoint (i.e. PONV-free patients) was the number of patients in the treatment arm who needed no treatment (T1).

In the prophylaxis arm both endpoints were the number of patients who never experienced any symptoms of PONV, either because they would have had none anyway or because prophylaxis was successful *(Figure 54:* P1 + P2).

Subgroup analyses

Subgroup analyses were performed to test the impact of different risks of PONV and of different doses of ondansetron on cost-effectiveness. It was assumed that the CER (i.e. what happens without anti-emetic prophylaxis) would accurately reflect the true underlying risk in a study population. Hence, two arbitrarily defined clinical settings were compared: a low-risk setting (CER, 30%) and a high-risk setting (CER, 60%).

Results

Modelling: graphical display

Number of patients who are PONV-free

For each CER value between 10% and 90%, prophylaxis yielded more PONV-free patients than treatment *(Figure 55A)*. This relationship was linear because it was assumed that ondansetron's anti-emetic efficacy was independent of CER. With a CER of 30% (low-risk setting), 700 of

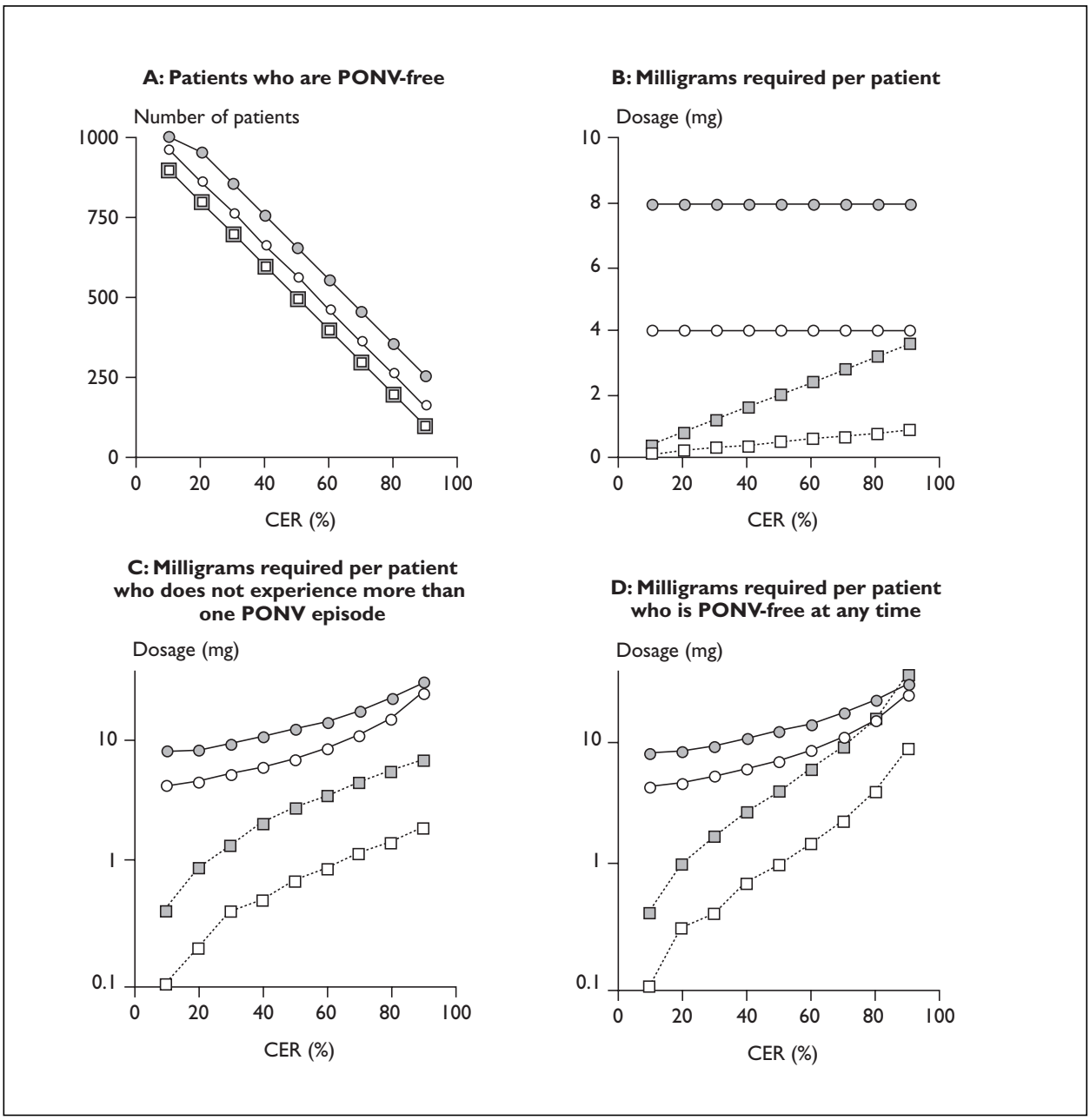


FIGURE 55 *PONV strategy – treatment versus prophylaxis (Treatment: □, 1 mg; ■, 4 mg. Prophylaxis: ○, 4 mg; ●, 8 mg)*

1000 patients will be completely PONV-free with treatment doses of both 1 mg and 4 mg, compared with 763 patients (+ 9%) with 4 mg prophylaxis and 856 patients (+ 22%) with 8 mg prophylaxis. With a CER of 60% (high-risk setting), 400 of 1000 patients will be completely PONV-free with treatment doses of both 1 mg and 4 mg, compared with 463 patients (+ 16%) with 4 mg prophylaxis and 556 patients (+ 39%) with 8 mg prophylaxis.

Cost per patient

The cost per patient, expressed as milligrams required per patient, was stable across all CERs with prophylaxis. With treatment, the cost per patient was related directly to CER value *(Figure 55B)*. Fewer milligrams per patient were required for each CER with 1 mg treatment than with 4 mg treatment. The highest cost per patient was with the most effective prophylactic dose, 8 mg.

Cost per patient who experiences at most one episode of PONV

Costs (milligrams required per patient) increased with increasing CERs with all strategies but were lowest with 1 mg and 4 mg treatment, respectively *(Figure 55C)*. For this endpoint, treatment with both doses remained more cost-effective than prophylaxis for all CERs tested.

Cost per patient who is PONV-free at any time

Fewest milligrams per PONV-free patient were required with 1 mg treatment *(Figure 55D)*. Treatment with 4 mg doses also demonstrated a lower cost per PONV-free patients than both prophylaxis doses but only below a CER of 80%.

Subgroup analyses

Treatment with 4 mg versus prophylaxis with 4 mg

In both the low-risk (CER = 30%) and high-risk setting (CER = 60%), 63 extra PONV-free patients (6% of all patients) were gained with 4 mg prophylaxis compared with 4 mg treatment *(Table 62A* and *B)*. However, in the low-risk setting, 3.3 times as many milligrams were required with prophylaxis than with treatment to achieve this *(Table 62A)*. In the high-risk setting it was 1.7 times more *(Table 62B)*.

Treatment with 1 mg versus prophylaxis with 8 mg

When 1 mg for treatment and 8 mg for prophylaxis were used, 156 extra PONV-free patients (16% of all patients) were gained with prophylaxis compared with treatment in both low- and high-risk settings *(Table 63A* and *B)*. However, for the low-risk setting 27 times as many milligrams were required with prophylaxis than with treatment to achieve this *(Table 63A)*. In the high-risk setting it was 13 times more *(Table 63B)*.

Incremental cost-effectiveness analysis

No more than one episode of PONV

With a low CER value (i.e. 30%), prophylaxis with 8 mg yielded the highest number of patients who experienced no more than one episode of PONV, while prophylaxis with 4 mg yielded the lowest number *(Table 64A)*. The crude calculation of the number of milligrams of ondansetron required to obtain one patient who experienced no more than one episode of PONV showed that the best result was treatment with 1 mg, at 0.4 mg/patient.

Changing from the least successful strategy, prophylaxis with a dose of 4 mg, to treatment with a dose of 1 mg resulted in a higher success rate and, because many fewer milligrams were required, resulted in a sparing effect: almost 65 mg were saved for each additional success. Only a slight improvement would be achieved by switching from 1 mg to 4 mg treatment; 60 mg would be required to generate one additional patient who had no more than one episode of PONV. Finally, another 324 mg would be required to gain one additional patient who profited when changing the strategy from treatment with 4 mg to prophylaxis with 8 mg.

With a high CER value (i.e. 60%), both treatment doses yielded more patients who had no more than one episode of PONV than both prophylaxis doses *(Table 64A)*. The crude calculation of milligrams of ondansetron required to obtain one patient who experienced no more than one episode of PONV showed that the best result was treatment with 1 mg, at 0.9 mg/patient.

Changing from prophylaxis at 4 mg to prophylaxis at 8 mg would result in an additional 93 patients who would profit, but this would require 43 mg per additional patient. Switching then to 1 mg treatment increases the number of patients who have no more than one episode of PONV and, at the same time, would save 88 mg for each additional case. Another 60 mg would be required per case when changing to 4 mg treatment; the gain would be a further 30 patients who experienced no more than one episode of PONV.

Patients who are PONV-free

With both low and high CERs, treatment with 1 mg or 4 mg yielded fewer patients who were PONV-free at any time compared with prophylaxis with 4 mg or 8 mg *(Table 64B)*. The crude calculation of milligrams of ondansetron required to obtain one patient who experienced no more than one

TABLE 62 *Treatment with ondansetron, 4 mg, versus prophylaxis with ondansetron, 4 mg*

Patients	Decision tree	CER	mg	NNT	No. of patients	Total mg spent	mg spent per success	mg spent per non-failure	mg spent per never PONV	mg spent per patient
A: CER = 30%										
Treatment of established PONV with ondansetron,4 mg										
No PONV (no need for treatment)	T1				700	0				
Successful treatment (45%)	T2				135	540				
Failure with treatment	T3				165	660				
All successes	T2				135	540	8.9			
All non-failures	T1 + T2				**835**	1200		1.4		
PONV-free	T1				**700**	0			1.7	
All patients	T1 + T2 + T3	30%	4		1000	**1200**				1.2
Prevention of PONV with ondansetron, 4 mg										
No PONV anyway	P1				700	2800				
No PONV because of prophylaxis (NNT 16)	P2				63	250				
PONV despite prophylaxis	P3				238	950				
All successes	P2				63	250	16.0			
All non-failures = PONV-free	P1 + P2				**763**	3050		5.2		
All patients	P1 + P2 + P3	30%	4	16.0	1000	**4000**				4.0
Difference in mg spent (factor x)						**3 x**				
Incremental cost-effectiveness ratio										
Difference in non-failures					–73	1850				
Difference in PONV-free					63	3050				
Difference in costs between treatment and prophylaxis						2800				
mg to generate an additional non-failure						**–26 mg**				
mg to generate an additional PONV-free						**49 mg**				
B CER = 60%										
Treatment of established PONV with ondansetron, 4 mg										
No PONV (no need for treatment)	T1				400	0				
Successful treatment (45%)	T2				270	1080				
Failure with treatment	T3				330	1320				
All successes	T2				270	1080	8.9			
All non-failures	T1 + T2				**670**	2400		3.6		
PONV-free	T1				**400**	0			6.0	
All patients	T1 + T2 + T3	60%	4		1000	**2400**				2.4
Prevention of PONV with ondansetron, 4 mg										
No PONV anyway	P1				400	1600				
No PONV because of prophylaxis (NNT 16)	P2				63	250				
PONV despite prophylaxis	P3				538	2150				
All successes	P2				63	250	16.0			
All non-failures = PONV-free	P1 + P2				**463**	1850		8.6		
All patients	P1 + P2 + P3	60%	4	16.0	1000	4000				4.0
Difference in mg spent (factor x)						**2 x**				
Incremental cost-effectiveness ratio										
Difference in non-failures					–208	–550				
Difference in PONV-free					63	1850				
Difference in costs between treatment and prophylaxis						1600				
mg to generate an additional non-failure						**3 mg**				
mg to generate an additional PONV-free						**30 mg**				

episode of PONV showed that the lowest result was treatment with 1 mg, at 0.4 mg/patient at a 30% CER and 1.5 mg/patient at a 60% CER.

However, switching from 1 mg or 4 mg treatment to prophylaxis with 4 mg would require 44 mg for each additional PONV-free patient when the CER was 30%, and 25 mg when the CER was 60%. A further change to 8 mg prophylaxis, in order to achieve the highest possible number of absolutely PONV-free patients, would require an additional 43 mg for each patients who benefited.

Drug-related adverse effects

With 4 mg or 8 mg prophylaxis, 33 extra patients from 1000 would have an adverse drug reaction who would not have had this reaction had they received placebo. Because all patients received the drug with prophylaxis, this outcome was independent of the CER. With treatment (1 mg or 4 mg),

TABLE 63 *Treatment with ondansetron, 1 mg, versus prophylaxis with ondansetron, 8 mg*

Patients	Decision tree	CER	mg	NNT	No. of patients	Total mg spent	mg spent per success	mg spent per non-failure	mg spent per never PONV	mg spent per patient
A CER = 30%										
Treatment of established PONV with ondansetron, 1 mg										
No PONV (no need for treatment)	T1				700	0				
Successful treatment (40%)	T2				120	120				
Failure with treatment	T3				180	180				
All successes	T2				120	120	2.5			
All non-failures	T1 + T2				**820**	300		0.4		
PONV-free	T1				**700**	0			0.4	
All patients	T1 + T2 + T3	30%	1		1000	**300**				0.3
Prevention of PONV with ondansetron, 8 mg										
No PONV anyway	P1				700	5600				
No PONV because of prophylaxis (NNT 6.4)	P2				156	1250				
PONV despite prophylaxis	P3				144	1150				
All successes	P2				156	1250	6.4			
All non-failures = PONV-free	P1 + P2				**856**	6850		9.3		
All patients	P1 + P2 + P3	30%	8	6.4	1000	**8000**				8.0
Difference in mg spent (factor x)						**27 x**				
Incremental cost effectiveness ratio										
Difference in non-failures					36	6550				
Difference in PONV-free					156	6850				
Difference in costs between treatment and prophylaxis						7700				
mg to generate an additional non-failure						**181 mg**				
mg to generate an additional PONV-free						**42 mg**				
B Control event rate 60%										
Treatment of established PONV with ondansetron, 1 mg										
No PONV (no need for treatment)	T1				400	0				
Successful treatment (40%)	T2				240	240				
Failure with treatment	T3				360	360				
All successes	T2				240	240	2.5			
All non-failures	T1 + T2				**640**	600		0.9		
PONV-free	T1				**400**	0			1.5	
All patients	T1 + T2 + T3	60%	1		1000	**600**				0.6
Prevention of PONV with ondansetron, 8 mg										
No PONV anyway	P1				400	3200				
No PONV because of prophylaxis (NNT 6.4)	P2				156	1250				
PONV despite prophylaxis	P3				444	3550				
All successes	P2				156	1250	6.4			
All non-failures = PONV-free	P1 + P2				**556**	4450		14.4		
All patients	P1 + P2 + P3	60%	8	6.4	1000	**8000**				8.0
Difference in mg spent (factor x)						**13 x**				
Incremental cost-effectiveness ratio										
Difference in non-failures					–84	3850				
Difference in PONV-free					156	4450				
Difference in costs between treatment and prophylaxis						7400				
mg to generate an additional non-failure						**–46 mg**				
mg to generate an additional PONV-free						**25 mg**				

ten extra patients experienced an adverse drug reaction when the CER was 30%, and 20 had such a reaction when the CER was 60%.

Comment

These results show that treatment of established PONV with ondansetron, 1 mg, is the most cost-effective option, irrespective of the proportion of patients who are likely to experience PONV. The systematic review suggests that ondansetron, 1 mg, is the most effective treatment and ondansetron, 8 mg, is the most effective prophylaxis. In practice, 4 mg is the recommended dose for both treatment and prophylaxis.

Using only the point estimates of effectiveness from systematic reviews, there was little difference in the number of patients who were either PONV-free or

TABLE 64 *Incremental costs to generate one additional patient who either experiences no more than one episode of PONV or is completely PONV-free*

A Not more than one episode of PONV							
Strategy	**Number with ≤ 1 episode PONV**	**Additional patients spared ≤ 1 episode PONV**	**Total mg spent**	**Additional mg required**	**'Cost' of generating one additional patient with ≤ 1 episode PONV (= x drug cost)**	**Comment**	**Total cost/ patients spared**
CER = 30%							
Prevention, 4 mg	763	N/A	4000	N/A	5		5.2
Treatment, 1 mg	820	57	300	–3700	–65	Sparing effect	0.4
Treatment, 4 mg	835	15	1200	900	60		1.4
Prevention, 8 mg	856	21	8000	6800	324		9.3
CER = 60%							
Prevention, 4 mg	463	N/A	4000	N/A	9		8.6
Prevention, 8 mg	556	93	8000	4000	43		14.4
Treatment, 1 mg	640	84	600	–7400	–88	Sparing effect	0.9
Treatment, 4 mg	670	30	2400	1800	60		3.6
B Completely PONV-free							
Strategy	**Number patients PONV-free**	**Additional PONV-free patients**	**Total mg required**	**Additional mg spent**	**'Cost' of generating one additional PONV-free patient (= x drug mg cost)**	**Comment**	**Total cost/ patients spared**
CER = 30%							
Treatment, 1 mg	700	N/A	300	N/A	0.4		0.4
Treatment, 4 mg	700	0	1200	900	> 900	No benefit	1.7
Prevention, 4 mg	763	63	4000	2800	44.4		5.2
Prevention, 8 mg	856	93	8000	4000	43.0		9.3
CER = 60%							
Treatment, 1 mg	400	N/A	600	N/A	1.5		1.5
Treatment, 4 mg	400	0	2400	1800	> 1800	No benefit	6.0
Prevention, 4 mg	463	63	4000	1600	25.4		8.6
Prevention, 8 mg	556	93	8000	4000	43.0		14.4

who experienced no more than one episode of PONV in any of these strategies. Because of this, the incremental analysis was highly unfavourable to prophylaxis because prophylaxis will consume very much higher quantities of anti-emetic but few additional patients will have a beneficial outcome. Even this may overstate the case, because point estimates have CIs about them. It might be considered legitimate to say that the number of additional patients apparently obtaining a benefit with prophylaxis was within the uncertainties of our estimates, so that no additional patients would be likely to benefit. If this were the case, prophylaxis exposes patients to much higher doses of drug for no benefit.

In money terms, how much more expensive prophylaxis will be depends on the unit cost of drug bought by hospital pharmacies. In practice this may be very much lower than the advertised price to GPs. However, hospitals in which prophylaxis with ondansetron was the norm would enjoy significant savings if a change were made to treatment of established PONV.

What might change this view? Unplanned overnight stays in day-case surgery because of PONV are clearly expensive but there is no reason to expect that this would be rarer with prophylactic ondansetron than with treatment. Adverse events consequent on the much greater use of ondansetron might incur greater costs – either through treatment of headache and an unplanned overnight stay because of headache, or through the unnecessary investigation of raised liver enzymes. Again, there is no evidence that these are major concerns.

Can these results be extrapolated to other anti-emetics? No direct comparisons have been found to indicate that any of the commonly prescribed anti-emetics are significantly better or worse than ondansetron (unpublished observations). Most other anti-emetics have a much lower cost per dose than ondansetron, and so monetary differences between prophylaxis and treatment may not be great. But if other anti-emetics do not differ in efficacy, but are cheaper, then that does point to other possible savings, albeit recognising the need also to examine their safety.

What is clear is that the great variability in results shown in individual trials, both in respect to PONV rates without treatment and in the apparent effect

of treatment,[1,2] makes economic evaluation pointless outside of evidence systematically gathered.

References

1. Tramèr MR, Moore RA, Reynolds DJM, McQuay HJ. A quantitative systematic review of ondansetron in treatment of established postoperative nausea and vomiting. *BMJ* 1997;**314**:1088–92.

2. Tramèr MR, Reynolds DJ, Moore RA, McQuay HJ. Efficacy, dose-response and safety of ondansetron in prevention of postoperative nausea and vomiting: a quantitative systematic review of randomised placebo-controlled trials. *Anesthesiology* 1997;**87**:1277–89.

3. Karlsson G, Johannesson M. The decision rules of cost-effectiveness analysis. *PharmacoEconomics* 1996;**9**:113–20.

Chapter 23

Conclusion

Interventions

In treating acute pain, as in other areas of medicine, tradition and ill-informed prejudice sometimes hold sway over evidence and common sense. This study concentrated on gathering evidence for the treatments which are simple. In this chapter that evidence is drawn together in a wider frame, including interventions for which there are no systematic reviews. Wherever possible our recommendations are based on randomised trials.

Effective pain management is fundamental to quality care and, while we consider that good pain control speeds recovery, there is still no compelling evidence that this is so, although advantage can be shown with proxy measures like mobility or coughing. However, evidence that good pain management led to faster recovery would increase the pressure to improve current practice, which is often less than ideal.

Non-opioids: paracetamol, combinations and NSAIDs

Effective relief can be achieved with oral non-opioids and NSAIDs. These drugs are appropriate for many post-surgical and post-traumatic pains, especially when patients go home on the day of the operation. The evolving league table for analgesic efficacy compiled from randomised trials after all kinds of surgery is shown in *Figure 56*. Analgesic efficacy is expressed as the NNT, the number of patients who need to receive the active drug for one to achieve at least 50% relief of pain compared with placebo over a 6-hour treatment period. The most effective drugs have low NNT values of about

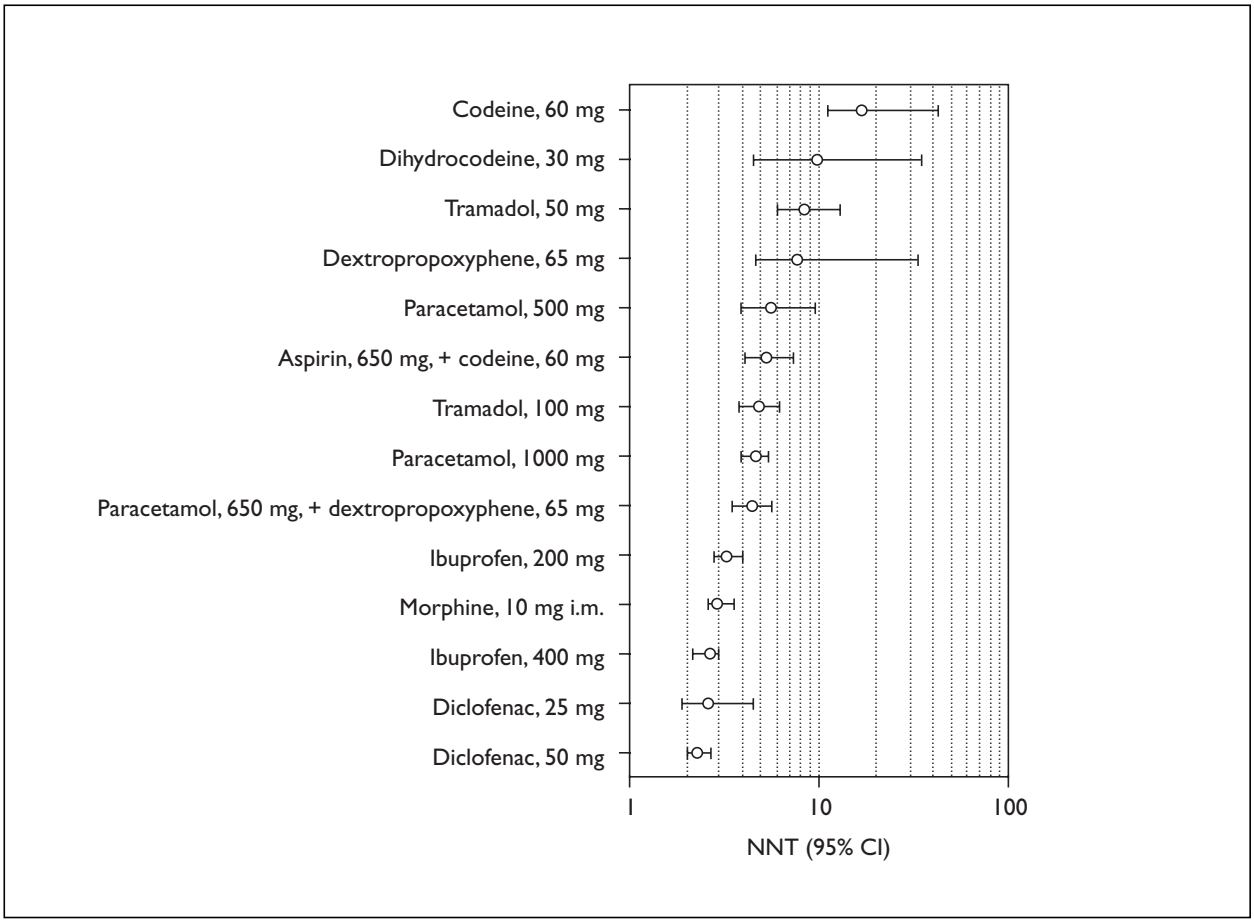


FIGURE 56 *Oral analgesic NNT league table*

2, meaning that for every two patients who receive the drug one patient will get at least 50% relief because of the treatment (the other patient may obtain relief but it does not reach the 50% level).

For paracetamol, 1 g, the NNT is nearly 5. Combination of paracetamol, 650 mg, with dextropropoxyphene, 65 mg, improves the NNT value slightly. Ibuprofen is better with an NNT of 3, as is diclofenac at about 2.5.

These NNT comparisons are against placebo; the best NNT value of 2 means that while 50 from 100 patients will get at least 50% relief because of the treatment, another 20% will have a placebo response which gives them at least 50% relief. Hence, with diclofenac, 70 from 100 patients will have effective pain relief.

This alternative way of looking at the effect of the various analgesics is shown in *Figure 57*. The range is from about 25% of patients getting at least 50% pain relief with codeine, 60 mg (largely because of the effect of placebo), to over 70%, at the high end of the 95% CI, with oral NSAID. For comparison, with morphine, 10 mg i.m., about 53% of patients get more than 50% pain relief. Because the effect of placebo is added in, the comparisons between analgesics are not as stark as with NNT.

The clear message is that, of the oral analgesics, NSAIDs perform best and paracetamol alone or in combination is also effective. Initial prescription of oral NSAIDs can be supplemented with paracetamol. As pain wanes then the prescription should be paracetamol-based, supplemented if necessary by NSAIDs.

There is an old adage that if a patient can swallow then it is best to take drugs by mouth. There is no evidence that NSAIDs given rectally or by injection perform better (or faster) than the same drug at the same dose given by mouth (see chapter 11). These other routes become appropriate when patients cannot swallow. Topical NSAIDs are effective in acute musculoskeletal injuries ibuprofen has an NNT of 3 for at least 50% relief at one week compared with placebo (see chapter 12).

Adverse effect data on NSAIDs from long-term dosing, where gastric bleeding is the main worry, rates ibuprofen the safest.[1] In acute pain, the main concerns are renal and coagulation problems. Acute renal failure can be precipitated in patients with pre-existing heart or kidney disease, those on loop diuretics or those who have lost more than 10% of blood volume. NSAIDs cause significant lengthening (~30%) of bleeding time, usually still within the normal range. This can last for days with aspirin, hours with non-aspirin-based NSAIDs. Whether or not NSAIDs cause significant increase in blood loss remains contentious.

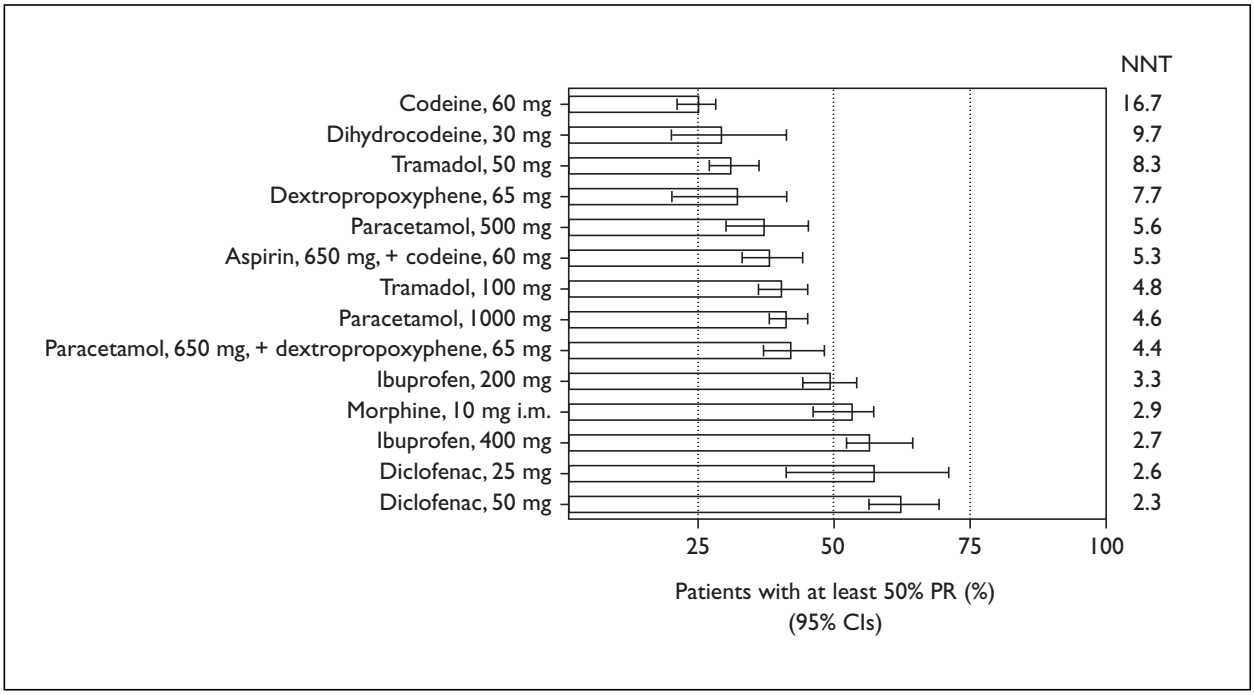


FIGURE 57 *Oral analgesic league table – percentage with at least 50% pain relief*

Other drugs

As yet there is no evidence from systematic reviews for a number of niche analgesic interventions. These include: inhaled nitrous oxide, which can provide fast-onset, fast-offset analgesia for obstetrics and wound dressings; corticosteroids to reduce pain and swelling after head and neck surgery, and when swelling causes pain in cancer; ketamine in emergency analgesia and anaesthesia; and clonidine.

Opioids

For severe acute pain, opioids are the first-line treatment and, to date, only one systematic review has been found – that for injected morphine (see chapter 13). Intermittent opioid injection can provide effective relief of acute pain.[2] Unfortunately, adequate doses are withheld because of traditions, misconceptions, ignorance and fear. Doctors and nurses fear addiction and respiratory depression. Addiction is not a problem with opioid use in acute pain. Over 11,000 patients were followed-up a year after opioids were given for acute pain, and just four were considered addicts.[3]

Irrespective of the route, opioids used for people who are not in pain, or in doses larger than necessary to control the pain, can slow or indeed stop breathing. The key principle is to titrate the dose against the desired effect – pain relief – and minimise unwanted effects *(Figure 58)*. If the patient is still complaining of pain and you are sure that the drug has all been delivered and absorbed, then it is safe to give another, usually smaller, dose (5 minutes after intravenous, 1 hour after intramuscular or subcutaneous, 90 minutes after oral). If the second dose is also ineffective, then the process should be repeated or the route of administration changed to achieve faster control. Delayed release formulations, oral or transdermal, should not be used in acute pain, because delayed onset and offset are dangerous in this context.

There is no compelling evidence that one opioid is better than another, but there is good evidence that pethidine has a specific disadvantage[4] and no specific advantage. Given in multiple doses, the metabolite norpethidine can accumulate and act as a central nervous system irritant, ultimately causing convulsions, especially in renal dysfunction. Pethidine should not be used when multiple injections are needed. The long-held view that pethidine is better than other opioids when dealing with colicky pain is no longer tenable.[5]

Morphine (and its relatives diamorphine and codeine) has an active rather than a toxic metabolite, morphine-6-glucuronide. In renal dysfunction, this metabolite can accumulate and result in greater effect from a given dose, because it is more active than morphine. If dose is being titrated against effect, as it should be, this will not matter. Less morphine will be needed. Accumulation can be a problem with unconscious intensive care patients on fixed dose schedules when renal function is compromised.

Opioid adverse effects include nausea and vomiting, constipation, sedation, pruritus, urinary retention and respiratory depression. There is no good evidence that the incidence is different with different opioids at the same level of analgesia. There is good evidence that the risk of adverse events is increased when high-technology approaches are used for drug administration.[6]

Principle for safe and effective opiate use – titrate to effect

If the patient is asking for more opioid than it usually signals inadequate pain control:

– too little drug
– too long between doses
– too little attention paid to the patient
– too much reliance on rigid (inadequate) prescriptions

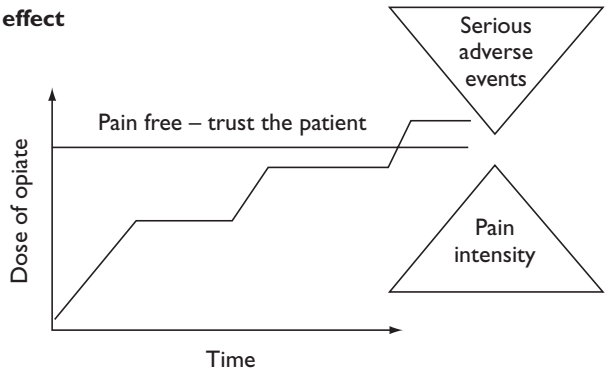


FIGURE 58 *Titrate opioids to effect*

There are strong arguments, based on minimising risk, for using one opioid only, so that everyone is familiar with dosage, effects and problems. Our first choice opioid is morphine. Whichever drug is chosen, simple changes in the way opioids are used, good staff education and implemention of an algorithm for intermittent opioid dosing, can have a powerful impact on pain relief and patient satisfaction.[2]

Nurse-administered intermittent opioid injection requires good staffing levels to minimise delay between need and injection. Staffing shortages, ward distractions and controlled drug regulations all increase the delay. PCA overcomes these logistical problems. The patient presses a button and receives a pre-set dose of opioid, from a syringe driver connected to an intravenous or subcutaneous cannula. This delivers the opioid to the same opioid receptors as an intermittent injection but allows the patient to circumvent delays. Not surprisingly, there is little difference in outcome between efficient intermittent injection and PCA.[7] Good risk management with PCA should emphasise the same drug, protocols and equipment throughout a hospital.

Novel routes of opioid administration, intended to improve analgesia and reduce adverse effects, include intra-articular (see chapter 17), nasal, active transdermal and inhalational. These may prove to have advantage over conventional routes, different kinetic profiles or greater convenience, but their place in mainstream care is unproven.

Regional analgesia

The perceived advantage of regional analgesia with local anaesthetic is that it can deliver complete pain relief by interrupting pain transmission from a localised area, so avoiding generalised drug adverse effects. This advantage is more obvious when it is possible to give further doses via a catheter, extending the duration of analgesia. Details are given in *Table 65*.

There is a necessary distinction between blocks undertaken to permit surgery, and blocks undertaken together with a general anaesthetic to provide postoperative pain relief. There is clear evidence that blocks can indeed provide good relief in the initial postoperative period[8] and no evidence to suggest that patients with blocks then experience 'rebound', and need more postoperative pain relief. The risk of neurological damage is the major drawback[9] and, ideally, blocks should not be undertaken on anaesthetised patients.

Epidural analgesia

Epidural infusion via a catheter can offer continuous relief after trauma or surgery for lower limb, spine, abdominal or chest. The current optimal infusate is an opioid/local anaesthetic mixture. Opioids and local anaesthetics have a synergistic effect, so that lower doses of each are required for equivalent analgesia with fewer adverse effects.[10]

The risks are those of an epidural (dural puncture, infection, haematoma, nerve damage), those of the local anaesthetic (hypotension, motor block, toxicity), and those of the opioid (nausea, sedation, urinary retention, respiratory depression, pruritus). Wrong doses do get administered,[6] so increased surveillance is mandatory. The risk of persistent neurological sequelae after an epidural is about 1 in 5000.[11] Debate continues as to whether patients with epidural infusions can be nursed on general wards. These techniques are only appropriate for major trauma or surgery when the potential benefits outweigh the risks.

TABLE 65 *Regional analgesia summary*

		Indications	Advantages	Problems
Low technology	Topical	Surface wounds	Simple	Short duration
	Wound infiltration	Most wounds	Simple	Short duration
	Peripheral nerve blocks	Limb surgery/trauma	Catheter possible	??
	Plexus blocks	Limb surgery	Catheter possible	Nerve damage, motor block
High technology	Epidural (including caudal)	Major surgery (thoracoabdominal, lower limb)	Catheter possible; reduced thrombo-embolism	Adverse effects surveillance
	Intrathecal	Major surgery (thoracoabdominal, lower limb)	Long duration relief possible from single injection low-dose opioid	

Other techniques

While experts can obtain good results with specialised procedures, such as paravertebral or interpleural injections, the evidence that in less skilled hands these are better than standard methods (should-do rather than can-do evidence) is often lacking.

TENS and acupuncture

TENS is not effective for postoperative pain (see chapter 20) and is of limited value for labour pain (see chapter 21). Systematic reviews of acupuncture are confined to chronic pain.

Psychological methods

There is evidence that psychological approaches are beneficial.[12] Cognitive behavioural methods can reduce pain and distress in patients with burns. Preparation before surgery can reduce postoperative analgesic consumption.

Clinical settings and recommendations (see *Box 4*)

General

The tenets of good management of acute pain are that, with good staff (and patient) education in place, appropriate drug doses are given when needed by the appropriate route and delivery method. Schemes have to be flexible enough to respond to individual patient need and different clinical settings. A general strategy is presented in *Figure 59*.

There is controversy about the optimal timing of initial analgesia. The idea is that analgesia is more effective when given before pain begins than when given after. Most randomised trials comparing the

BOX 4 How to achieve successful pain management

Factors to consider when choosing therapy
Co-existing illness
Staff availability
Equipment available
Risks and unwanted effects of the various options
Appropriateness of the chosen intervention for that pain
Evidence of efficacy for the chosen intervention
Cost

Steps to successful management
Regular assessment of pain and adverse effects
Protocols for monitoring and treating pain
Protocols for monitoring and treating adverse effects
Titration of doses at short intervals until pain relieved
Consideration of more than one approach
Appropriate back-up by identified personnel
Continuing in-service training and education

Predictable problems

Patient	Problems
Babies and infants	Communication; drug handling
Elderly	Co-existing illness; drug handling
Respiratory disease	Respiratory depression; NSAIDs and asthma
Renal failure	Drug handling; NSAIDs
Head injury or impaired consciousness	Assessment; dose titration
Drug addiction or already taking opioids	Dose titration; weaning; respiratory depression after nerve block which stops pain
Sickle cell disease	Assessment; varying analgesic needs

Key points for improving acute pain management
1 Opt for safety and simplicity.
2 Measure and record pain regularly – be proactive.
3 Choose evidence-based interventions.
4 Individualise treatment and allow patient to control analgesia.
5 Choose appropriate drug, route and mode of delivery.
6 Provide education for staff and patients.

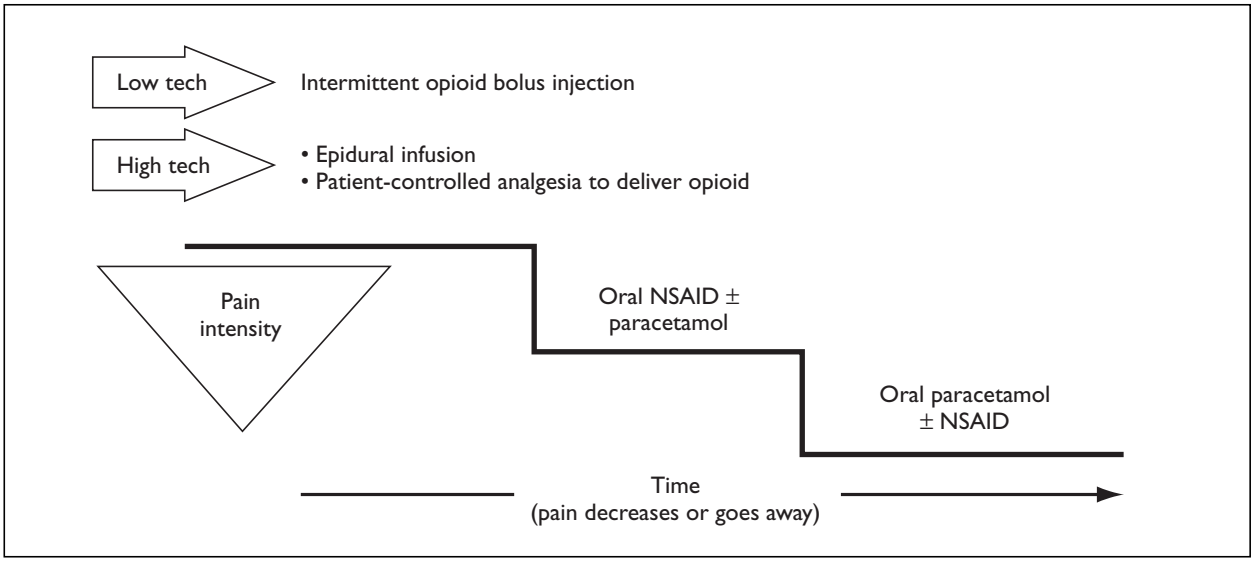


FIGURE 59 *General strategy*

same intervention given before or after pain starts have not shown clinical advantage of so-called pre-emptive analgesia.[13] Whether poorly-controlled acute pain generates chronic pain is also controversial.

Problem pains and patients

Standard interventions and protocols will cope with most acute pain problems but some patients, particularly in hospital, will require special management. Expertise can be developed in specific units but, if not available, the advice of your acute pain service should be sought. In particular do not let pain go untreated in children.

Conclusion

The key to successful pain management is education, not new drugs or high-technology delivery systems. Existing tools can do the job if doctors and nurses are educated, both to dispel the myths and misconceptions and to take responsibility for providing good pain control. It is much easier to dispel myths when you have the evidence. For many years patients were not given adequate analgesia for abdominal pain in case it masked the signs necessary for diagnosis. This was wrong.[14]

Pain relief should not be seen as someone else's responsibility; nor should it be just dismissed, because 'in the end the pain and the patient go away'. Freedom from pain is important to patients. In 1846, the first anaesthetic provided pain-free surgery. Some 150 years later patients should not have to endure unrelieved pain anywhere in hospital.

References

1. Henry D, Lim LL, Rodriguez LAG, Gutthann SP, Carson JL, Griffin M, *et al.* Variability in risk of gastrointestinal complications with individual non-steroidal anti-inflammatory drugs: results of a collaborative meta-analysis. *BMJ* 1996;**312**:1563–6.
2. Gould TH, Crosby DL, Harmer M, Lloyd SM, Lunn JN, Rees GAD, *et al.* Policy for controlling pain after surgery: effect of sequential changes in management. *BMJ* 1992;**305**:1187–93.
3. Porter J, Jick H. Addiction rate in patients treated with narcotics. *N Engl J Med* 1980;**302**:123.
4. Szeto HH, Inturrisi CE, Houde R, Saal S, Cheigh J, Reidenberg M. Accumulation of norperidine, an active metabolite of meperidine, in patients with renal failure or cancer. *Ann Intern Med* 1977;**86**:738–41.
5. Nagle CJ, McQuay HJ. Opiate receptors; their role in effect and side-effect. *Curr Anaesth Crit Care* 1990;**1**:247–52.
6. Bates DW, Cullen DJ, Laird N, Petersen LA, Small SD, Servi D, *et al.* Incidence of adverse drug events and potential adverse drug events. *JAMA* 1995;**274**:29–34.
7. Ballantyne JC, Carr DB, Chalmers TC, Dear KB, Angelillo IF, Mosteller F. Postoperative patient-controlled analgesia: meta-analyses of initial randomized control trials. *J Clin Anesth* 1993;**5**:182–93.
8. McQuay HJ, Carroll D, Moore RA. Postoperative orthopaedic pain – the effect of opiate premedication and local anaesthetic blocks. *Pain* 1988;**33**:291–5.
9. Bridenbaugh PO. Complications of local anesthetic neural blockade. In: Cousins MJ, Bridenbaugh PO, editors. Neural blockade. 2nd edition. Philadelphia: Lippincott, 1988:695–717.
10. McQuay H. Epidural analgesics. In: Wall P, Melzack R, editors. Textbook of pain. 3rd edition. London: Churchill Livingstone, 1994:1025–34.
11. Kane RE. Neurologic deficits following epidural or spinal anesthesia. Anesth Analg 1981;60:150–61.
12. Justins DM, Richardson PH. Clinical management of acute pain. *Br Med Bull* 1991;**47**:561–83.
13. McQuay HJ. Pre-emptive analgesia: a systematic review of clinical studies. *Ann Med* 1995;**27**:249–56.
14. Attard AR, Corlett MJ, Kidner NJ, Leslie AP, Fraser IA. Safety of early pain relief for acute abdominal pain. *BMJ* 1992;**305**:554–6.

Acknowledgements

The authors owe a huge debt to our colleagues in Oxford and around the world who have contributed to this work. Alejandro Jadad, Dawn Carroll and our volunteers broke the back of hand-searching for randomised trials of pain-relieving interventions. The chapters on finding trials, on appraising their quality and on existing systematic reviews borrow heavily from the publications which came from Alejandro's DPhil thesis. The Oxford contributions of Dawn Carroll, Sally Collins, David Gavaghan, Jayne Edwards, John Reynolds and Phil Wiffen are gratefully acknowledged. Without their endeavour and their patience much of this would not have happened. Martin Tramèr (Oxford and Geneva), Eija Kalso (Helsinki), Pascale Picard (Clermont-Ferrand), Clara Faura (Alicante) and Göran Leijon (Linköping) collaborated on particular reviews. Martin applied the same obsession he used on systematic reviews of anti-emetics to reviews of analgesics, and his work on anti-emetics is summarised in later chapters, together with an economic analysis which includes much thought and work by Ceri Phillips, our health economics correspondent from Swansea. Our clinical colleagues Chris Glynn and Tim Jack have been helpful, enthusiastic and protective sounding-boards.

The authors' thinking over the past 6 years or so has been informed by discussion and argument with some very special people who were either in Oxford or who have come here to work. They include Muir Gray, Iain Chalmers, David Sackett, Doug Altman and Jon Deeks. Many other people have acted as sounding-boards, such as Douglas Justins, Peter Evans, Ed Charlton and Tony Dickenson. Others have helped and supported, including John Williams, Dick Waite, Ann Southwell and Aprille Cornell. Our libraries in and around Oxford have been magnificent, especially Claire Abbott here in the Churchill Hospital, as have many libraries in academia and industry who have provided the raw material for the work.

Financial support for the work has been provided from many different sources for different aspects of the work, including: Pain Research Funds, European Union Biomed 2, NHS Research and Development HTA Programme, and NHSE Anglia and Oxford Research and Development Directorate, as well as pharmaceutical companies.

Our thanks are also due to the referees for their perseverance in reading the report and for the quality of their comments.

Finally, it is a pleasure to acknowledge our old mentor, John Lloyd. Not only did John start the Oxford Pain Relief Unit but he set us and many others on a quest for better treatments and greater understanding.

Health Technology Assessment panel membership

This report was identified as a priority by the Pharmaceutical Panel.

Acute Sector Panel

Chair: Professor John Farndon, University of Bristol †

Professor Senga Bond, University of Newcastle-upon-Tyne †

Professor Ian Cameron, Southeast Thames Regional Health Authority

Ms Lynne Clemence, Mid-Kent Health Care Trust †

Professor Francis Creed, University of Manchester †

Professor Cam Donaldson, University of Aberdeen

Mr John Dunning, Papworth Hospital, Cambridge †

Professor Richard Ellis, St James's University Hospital, Leeds

Mr Leonard Fenwick, Freeman Group of Hospitals, Newcastle-upon-Tyne †

Professor David Field, Leicester Royal Infirmary †

Ms Grace Gibbs, West Middlesex University Hospital NHS Trust †

Dr Neville Goodman, Southmead Hospital Services Trust, Bristol †

Professor Mark P Haggard, MRC †

Mr Ian Hammond, Bedford & Shires Health & Care NHS Trust

Professor Adrian Harris, Churchill Hospital, Oxford

Professor Robert Hawkins, University of Bristol †

Dr Gwyneth Lewis, Department of Health †

Dr Chris McCall, General Practitioner, Dorset †

Professor Alan McGregor, St Thomas's Hospital, London

Mrs Wilma MacPherson, St Thomas's & Guy's Hospitals, London

Professor Jon Nicholl, University of Sheffield †

Professor John Norman, University of Southampton

Dr John Pounsford, Frenchay Hospital, Bristol †

Professor Gordon Stirrat, St Michael's Hospital, Bristol

Professor Michael Sheppard, Queen Elizabeth Hospital, Birmingham †

Dr William Tarnow-Mordi, University of Dundee

Professor Kenneth Taylor, Hammersmith Hospital, London

Diagnostics and Imaging Panel

Chair: Professor Mike Smith, University of Leeds †

Professor Michael Maisey, Guy's & St Thomas's Hospitals, London *

Professor Andrew Adam, UMDS, London †

Dr Pat Cooke, RDRD, Trent Regional Health Authority

Ms Julia Davison, St Bartholomew's Hospital, London †

Professor Adrian Dixon, University of Cambridge †

Mr Steve Ebdon-Jackson, Department of Health †

Professor MA Ferguson-Smith, University of Cambridge †

Dr Mansel Hacney, University of Manchester

Professor Sean Hilton, St George's Hospital Medical School, London

Mr John Hutton, MEDTAP International Inc., London

Professor Donald Jeffries, St Bartholomew's Hospital, London †

Dr Andrew Moore, Editor, *Bandolier* †

Professor Chris Price, London Hospital Medical School †

Dr Ian Reynolds, Nottingham Health Authority

Professor Colin Roberts, University of Wales College of Medicine

Miss Annette Sergeant, Chase Farm Hospital, Enfield

Professor John Stuart, University of Birmingham

Dr Ala Szczepura, University of Warwick †

Mr Stephen Thornton, Cambridge & Huntingdon Health Commission

Dr Gillian Vivian, Royal Cornwall Hospitals Trust †

Dr Jo Walsworth-Bell, South Staffordshire Health Authority †

Dr Greg Warner, General Practitioner, Hampshire †

Methodology Panel

Chair: Professor Martin Buxton, Brunel University †

Professor Anthony Culyer, University of York *

Dr Doug Altman, Institute of Health Sciences, Oxford †

Professor Michael Baum, Royal Marsden Hospital

Professor Nick Black, London School of Hygiene & Tropical Medicine †

Professor Ann Bowling, University College London Medical School †

Dr Rory Collins, University of Oxford

Professor George Davey-Smith, University of Bristol

Dr Vikki Entwistle, University of Aberdeen †

Professor Ray Fitzpatrick, University of Oxford †

Professor Stephen Frankel, University of Bristol

Dr Stephen Harrison, University of Leeds

Mr John Henderson, Department of Health †

Mr Philip Hewitson, Leeds FHSA

Professor Richard Lilford, Regional Director, R&D, West Midlands †

Mr Nick Mays, King's Fund, London †

Professor Ian Russell, University of York †

Professor David Sackett, Centre for Evidence Based Medicine, Oxford †

Dr Maurice Slevin, St Bartholomew's Hospital, London

Dr David Spiegelhalter, Institute of Public Health, Cambridge †

Professor Charles Warlow, Western General Hospital, Edinburgh †

* Previous Chair
† Current members

continued

continued

Pharmaceutical Panel

Chair: Professor Tom Walley, University of Liverpool †

Professor Michael Rawlins, University of Newcastle-upon-Tyne*

Dr Colin Bradley, University of Birmingham

Professor Alasdair Breckenridge, RDRD, Northwest Regional Health Authority

Ms Christine Clark, Hope Hospital, Salford †

Mrs Julie Dent, Ealing, Hammersmith & Hounslow Health Authority, London

Mr Barrie Dowdeswell, Royal Victoria Infirmary, Newcastle-upon-Tyne

Dr Tim Elliott, Department of Health †

Dr Desmond Fitzgerald, Mere, Bucklow Hill, Cheshire

Dr Felicity Gabbay, Transcrip Ltd †

Dr Alistair Gray, Health Economics Research Unit, University of Oxford †

Professor Keith Gull, University of Manchester

Dr Keith Jones, Medicines Control Agency

Professor Trevor Jones, ABPI, London †

Ms Sally Knight, Lister Hospital, Stevenage †

Dr Andrew Mortimore, Southampton & SW Hants Health Authority †

Mr Nigel Offen, Essex Rivers Healthcare, Colchester †

Dr John Posnett, University of York

Mrs Marianne Rigge, The College of Health, London †

Mr Simon Robbins, Camden & Islington Health Authority, London †

Dr Frances Rotblat, Medicines Control Agency †

Mrs Katrina Simister, Liverpool Health Authority †

Dr Ross Taylor, University of Aberdeen †

Dr Tim van Zwanenberg, Northern Regional Health Authority

Dr Kent Woods, RDRD, Trent RO, Sheffield †

Population Screening Panel

Chair: Professor Sir John Grimley Evans, Radcliffe Infirmary, Oxford †

Dr Sheila Adam, Department of Health*

Ms Stella Burnside, Altnagelvin Hospitals Trust, Londonderry †

Dr Carol Dezateux, Institute of Child Health, London †

Dr Anne Dixon Brown, NHS Executive, Anglia & Oxford †

Professor Dian Donnai, St Mary's Hospital, Manchester †

Dr Tom Fahey, University of Bristol †

Mrs Gillian Fletcher, National Childbirth Trust †

Professor George Freeman, Charing Cross & Westminster Medical School, London

Dr Mike Gill, Brent & Harrow Health Authority †

Dr JA Muir Gray, RDRD, Anglia & Oxford RO †

Dr Anne Ludbrook, University of Aberdeen †

Professor Alexander Markham, St James's University Hospital, Leeds †

Professor Theresa Marteau, UMDS, London

Dr Ann McPherson, General Practitioner, Oxford †

Professor Catherine Peckham, Institute of Child Health, London

Dr Connie Smith, Parkside NHS Trust, London

Dr Sarah Stewart-Brown, University of Oxford †

Ms Polly Toynbee, Journalist †

Professor Nick Wald, University of London †

Professor Ciaran Woodman, Centre for Cancer Epidemiology, Manchester

Primary and Community Care Panel

Chair: Dr John Tripp, Royal Devon & Exeter Healthcare NHS Trust †

Professor Angela Coulter, King's Fund, London *

Professor Martin Roland, University of Manchester *

Dr Simon Allison, University of Nottingham

Mr Kevin Barton, East London & City Health Authority †

Professor John Bond, University of Newcastle-upon-Tyne †

Ms Judith Brodie, Age Concern, London †

Dr Nicky Cullum, University of York †

Professor Shah Ebrahim, Royal Free Hospital, London

Mr Andrew Farmer, Institute of Health Sciences, Oxford †

Ms Cathy Gritzner, The King's Fund †

Professor Andrew Haines, RDRD, North Thames Regional Health Authority

Dr Nicholas Hicks, Oxfordshire Health Authority †

Professor Richard Hobbs, University of Birmingham †

Professor Allen Hutchinson, University of Sheffield †

Mr Edward Jones, Rochdale FHSA

Professor Roger Jones, UMDS, London

Mr Lionel Joyce, Chief Executive, Newcastle City Health NHS Trust

Professor Martin Knapp, London School of Economics & Political Science

Dr Phillip Leech, Department of Health †

Professor Karen Luker, University of Liverpool

Professor David Mant, NHS Executive South & West †

Dr Fiona Moss, Thames Postgraduate Medical and Dental Education †

Professor Dianne Newham, King's College London

Professor Gillian Parker, University of Leicester †

Dr Robert Peveler, University of Southampton †

Dr Mary Renfrew, University of Oxford

Ms Hilary Scott, Tower Hamlets Healthcare NHS Trust, London †

* Previous Chair
† Current members

Health Technology Assessment 1998; Vol. 2: No. 12

Copies of this report can be obtained from:

The National Coordinating Centre for Health Technology Assessment,
Mailpoint 728, Boldrewood,
University of Southampton,
Southampton, SO16 7PX, UK.
Fax: +44 (0) 1703 595 639 Email: hta@soton.ac.uk
http://www.soton.ac.uk/~hta

ISSN 1366-5278