

Systematic review of outpatient services for chronic pain control

Henry J McQuay

R Andrew Moore

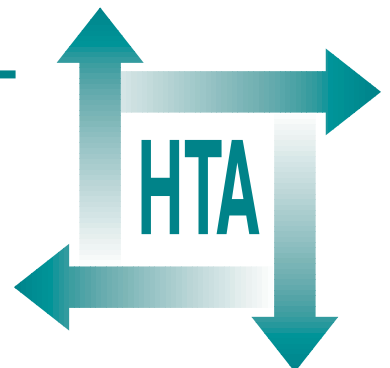
Christopher Eccleston

Stephen Morley

Amanda C de C Williams



**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,
Kings Fund Centre for Health Services
Development †

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 †

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

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

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

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

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

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 *

Mr Peter Bower,
Independent Management Consultant,
Newcastle-upon-Tyne †

Ms Christine Clarke, Director of
Pharmacy, Hope Hospital, Salford †

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

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

Dr Mike Gill, Brent and 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 Martin Knapp, Director,
Personal Social Services Research Unit,
London School of Economics and
Political Science

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

Professor Sally McIntyre, MRC Medical
Sociology Unit, Glasgow

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

Dr Tim Peters, Department of Social
Medicine, University of Bristol †

Professor David Sackett,
Centre for Evidence Based Medicine, Oxford

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

* Previous Chair
† Current members

National Coordinating Centre for Health Technology Assessment, Advisory Group

Chair: Professor John Gabbay, Wessex Institute for Health Research & Development

Professor Mike Drummond, Centre for
Health Economics, University of York

Professor Charles Florey,
Department of Epidemiology & Public Health,
University of Dundee

Professor Sir Miles Irving,
Department of Surgery, Hope Hospital, Salford

Ms Lynn Kerridge, Wessex Institute for
Health Research & Development

Dr Ruairidh Milne, Wessex Institute for
Health Research & Development

Ms Kay Pattison, Research &
Development Directorate, NHS Executive

Professor James Raftery, Health
Economics Unit, University of Birmingham

Dr Paul Roderick, Wessex Institute for
Health Research & Development

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

Dr Ken Stein, Wessex Institute for Health
Research & Development

Professor Andrew Stevens,
Department of Public Health &
Epidemiology, University of Birmingham



INAHTA

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.

Systematic review of outpatient services for chronic pain control

Henry J McQuay¹

R Andrew Moore¹

Christopher Eccleston^{2*}

Stephen Morley^{3*}

Amanda C de C Williams^{4*}

¹ Pain Research & Nuffield Department of Anaesthetics, University of Oxford, Oxford Radcliffe Hospital Trust, Churchill Hospital, Oxford

² University of Bath, Pain Management Unit, RNHRD, Bath

³ Division of Psychiatry & Behavioural Sciences, School of Medicine, University of Leeds

⁴ INPUT, Pain Management Unit, St Thomas' Hospital, London

* Authors of chapter 19, Psychological approaches, only

Final manuscript received April 1997

Published July 1997

This report should be referenced as follows:

McQuay HJ, Moore RA, Eccleston C, Morley S, de C Williams AC. Systematic review of outpatient services for chronic patient control. *Health Technol Assessment* 1997; **1**(6).

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 Acute Sector Panel (see inside back cover).

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.

Series Editors: Andrew Stevens, Ruairidh Milne and Ken Stein
Assistant Editor: Jane Robertson

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

© Crown copyright 1997

Enquiries relating to copyright should be addressed to the NCCHTA (see address given below).

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	i	TENS in labour pain	45
Executive summary	iii	Labour pain – results	45
1 Systematic reviews, efficacy and safety	1	TENS in chronic pain	46
Efficacy	1	Chronic pain – results	47
Safety	1	Comment	47
Application to chronic pain	2	Conclusion	49
2 Finding all the relevant trials	3	9 Relaxation for relief of chronic pain	51
Developing a citation database	3	10 Intravenous regional sympathetic blockade for reflex sympathetic dystrophy	53
Conclusions	5	Summary	53
3 Judging the quality of trials	7	Introduction	53
Developing and validating a quality scale	7	Systematic review	53
Comments on the scale	8	Comment	55
4 Pain measurement, study design and validity	11	11 Epidural corticosteroids for sciatica	57
Pain measurement	11	12 Spinal cord stimulation for back pain	61
Study design and validity	12	13 Steroid injections for shoulder disorders..	63
5 Estimating relative effectiveness	15	Comment	63
Placebo responses in analgesic trials	15	14 Anticonvulsant drugs	65
Deriving dichotomous outcome measures from continuous data in RCTs of analgesics ..	17	Summary	65
Verification from independent data	23	Introduction	65
Use of pain intensity and VAS	25	Methods	65
Comment	29	Results	66
Relative efficacy of minor analgesics	30	Discussion	72
6 Combining data and interpreting the results	31	15 Antidepressants in neuropathic pain	75
Combining data: quantitative systematic reviews	32	Summary	75
Statistical significance	34	Introduction	75
How well does the intervention work?	35	Methods	75
Is it safe?	37	Results	76
Using NNTs	37	Discussion	82
7 Existing systematic reviews	39	16 Systemic local anaesthetic type drugs in chronic pain	85
Introduction	39	Summary	85
Methods	39	Introduction	85
Results	40	Methods	85
Comments	42	Results	86
8 Transcutaneous electrical nerve stimulation	43	Comment	90
TENS in acute pain	43	17 Topical NSAIDs in chronic painful conditions	91
Acute pain – results	44	Methods	91
		Results	91

Discussion	92	Cost of chronic pain treatment	112
Comment	93	Comment	112
18 Topical capsaicin	95	21 Conclusions	115
Does it work?	95	Costs	115
Comment	96	Comment	115
19 Psychological approaches	97	Research recommendations	116
Introduction	97	Acknowledgements	117
Psychologically-based pain treatments	97	References	119
Reviews and analyses	98	Systematic reviews in chronic pain	130
Interim summary of averaged effect sizes over the studies by different outcome domains	103	Psychological reviews and RCTs	134
Model studies	107	Health Technology Assessment reports published to date	137
Conclusions	108		
20 Chronic pain services – evidence on costs	111		
Prevalence	111		



List of abbreviations

catPR	categorical verbal rating scale of pain relief
CBT	cognitive-behavioural therapy
CER	control event rate
CI	confidence interval
EER	experimental event rate
ESR	erythrocyte sedimentation rate
IRSB	intravenous regional sympathetic blockade
maxTOTPAR	maximum possible total pain relief (as a %)
NNH	number-needed-to-harm
NNT	number-needed-to-treat
NSAID	non-steroidal anti-inflammatory drug
PCA	patient-controlled analgesia
PHN	postherpetic neuralgia *
PID	pain intensity differences
RCT	randomised controlled trial
RWJ	Robert Wood Johnson Pharmaceutical Research Institute, Spring House, PA, USA
RSD	reflex sympathetic dystrophy
SCS	spinal cord stimulation
SD	standard deviation
SPID	sum of pain intensity differences
SSRI	selective serotonin reuptake inhibitor
TENS	transcutaneous electrical nerve stimulation
TN	trigeminal neuralgia *
TOTPAR	total pain relief
VAS	visual analogue scale

* Used only in figures and tables



Executive summary

Aim of report

This report reviews the evidence about the effectiveness of treatments for chronic pain. While treatment of chronic pain is usually seen as an integrated service, this report concentrates on the individual interventions that constitute the service.

How the research was conducted

Searches of databases and journals identified over 15,000 randomised studies with pain as an outcome, and many more which were not randomised. Over 150 systematic reviews relevant to chronic pain treatment were identified and their quality assessed using a simple scoring system. Systematic reviews conducted for this report were based mainly on randomised trials.

The number needed to treat (NNT) was chosen as the output for the report. NNTs of 2–4 indicate effective treatments. Because NNT is treatment-specific it overcomes problems associated with highly variable placebo or control event rates in pain trials. Such variability is predominantly due to the limited numbers of patients in the clinical trials.

Dichotomous outcome measures are important in synthesising information from many studies, and in deriving NNTs. Methods have been developed which allow mean information on pain relief and intensity to be converted reliably into the simple dichotomous outcome of at least 50% pain relief.

Research findings

Physical interventions

Transcutaneous electrical nerve stimulation (TENS) has been shown not to be effective in postoperative and labour pain. In chronic pain, there is evidence that TENS effectiveness increases slowly, and that large doses need to be used. There is lack of evidence for the effectiveness of TENS in chronic pain.

There is a lack of evidence for the effectiveness of **relaxation**.

Intravenous systemic regional blockade with guanethidine has been shown to be without effect.

Epidural corticosteroids are effective in the short term for back pain and sciatica.

Injections of corticosteroids in or around shoulder joints for shoulder pain have been shown not to be effective.

There is a lack of evidence supporting **spinal cord stimulators**. Case series are of poor quality and do not provide evidence of effectiveness, although at least 50% pain relief at 5 years is reported in over 50% of patients.

Pharmacological interventions

Minor analgesics are important in chronic pain. NNTs were calculated for analgesics given orally for moderate or severe acute postoperative pain. The NNTs found ranged from 17 (poor) for codeine, 60 mg, to 2.5 (good) for ibuprofen, 400 mg.

Anticonvulsant and **antidepressant drugs** are prescribed for neuropathic pains like diabetic neuropathy. NNTs are of the order of 2.5, showing them to be effective treatments. However, there are too few studies with too few patients to determine which is the best drug. Minor adverse events are common, and major adverse events occur in about 1 in 20 patients. There are no studies comparing antidepressants and anticonvulsants directly.

Systemic local anaesthetic-type drugs have been shown to be effective in nerve injury pain but there is little or no evidence to support their use in migraine or cancer-related pain.

Topical NSAIDs (for example, gels, creams) are effective in rheumatological conditions with an overall NNT of 3. There are too few studies to determine which is the best agent. Topical NSAIDs have few adverse events; most importantly they are without the major gastrointestinal adverse events found with oral NSAIDs, which might make them an important choice for some patients with peripheral arthritis.

In diabetic neuropathy, **topical capsaicin** has an NNT of 4, showing it to be effective, although the review contained no information about adverse events.

Psychological approaches

Cognitive-behavioural therapies provide strong evidence for efficacy across a range of mental health problems. Preliminary evidence from 35 trials in pain therapy demonstrates large and sustainable improvements in targeted outcomes.

Cost

While there is evidence that chronic pain clinics use interventions which provide pain relief for patients, there is little information on costs and benefits of chronic pain treatments. The evidence that is available suggests that pain clinics reduce overall direct healthcare costs by about £1000 per patient per year. The evidence indicates that pain clinics generate direct health service savings equal to twice their running cost.

Conclusions

The findings show that there is excellent evidence of effectiveness for some common treatments

for chronic pain, good evidence that some treatments are without effect, and a lack of evidence of effectiveness for some commonly-used treatments.

With regard to costing services, chronic pain units may save the National Health Service substantial sums by caring for patients and minimising unnecessary consultations and investigations. Given that there is substantial evidence for efficacy and inefficacy of individual interventions, the ideal would be for a process analysis approach to chronic pain services. This could well establish a model for other chronic services.

Research recommendations

- High quality randomised trials are needed in a number of different areas.
- The establishment of a single UK centre to organise and advise on large multicentre chronic pain studies may be appropriate.

Chapter I

Systematic reviews, efficacy and safety

There are both professional and political reasons for determining whether the interventions we use in health care are effective and safe. The professional agenda is that we want to use those interventions for our patients. The political agenda is that, with finite resources, it makes sense to pay for the effective and not to pay for the ineffective.

The problem is working out what is effective and what is safe. The tool used here is the systematic review. The term systematic review is used generically to encompass both qualitative reviews, in which no data-pooling was possible or none was done, and quantitative reviews, in which data-pooling (meta-analysis) was undertaken. Systematic reviews are different from classical narrative reviews because they have explicit methods, which describe the systematic way in which all the relevant studies have been identified and considered. Systematic reviews should be less open to bias than narrative reviews, and should be repeatable using the authors' methods – you might not come to the same conclusions, but at least you would be working from the same data.

This study is not designed as a 'cookbook' for chronic pain management. The structure is to provide the tools and the methods that are required to compare the efficacy of different interventions, and then to give examples of reviews of particular interventions. Without such efficacy data, neither patient nor professional nor the commissioner of healthcare can make rational decisions. In particular, cost-benefit analysis is irrational without good quality efficacy data.

Type and strength of evidence

- I Strong evidence from at least one systematic review of multiple well-designed RCTs
- II Strong evidence from at least one properly designed RCT of appropriate size
- III Evidence from well designed trials without randomisation, single group pre-post, cohort, time series or matched case-controlled studies
- IV Evidence from well-designed non-experimental studies from more than one centre or research group
- V Opinions of respected authorities, based on clinical evidence, descriptive studies or reports of expert committees.

FIGURE 1 Grading of efficacy evidence by study architecture

This review was prepared under a 1-year grant from the NHS Health Technology Assessment programme. Because the grant was for just 1 year, the review cannot be encyclopaedic about chronic pain. It does, however, provide a starting point for the acquisition of adequate knowledge about efficacy to permit subsequent (informed) investigations into the process and the cost-effectiveness of chronic pain management.

Efficacy

The rules for guiding us to the 'best' evidence about efficacy are relatively clear (see *Figure 1*). Opinions of experts (grade V) are less likely to supply the correct answer about efficacy than randomised trials or systematic reviews (grade I).

The arguments for restricting reviews to randomised controlled trials (RCTs) (if these are available) will be developed later. It is important to realise that even if there are many RCTs relevant to the question you want to pose, it may still not be possible for a systematic review to deliver an answer (*Figure 2*).

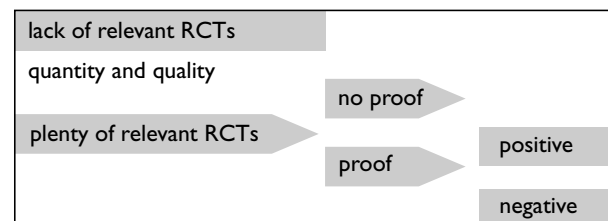


FIGURE 2 Possible outcomes when posing a question for systematic review

If all trials use different outcome measures, or if each trial is invalid for some bizarre methodological reason, then there may be no answer either way to your question. With plenty of relevant trials that do pass muster on quality and validity standards, there may be proof that an intervention is either effective or ineffective.

Safety

The rules for the quality of evidence about safety are not as well developed as those for efficacy. Whereas a case-report about efficacy should carry very little

weight, because of the strong possibility of bias, a case-report of a serious adverse event occurring in a 'benign' setting – death after tonsillectomy, for instance – may be extremely important. Rare events are unlikely to be identified in the relatively small numbers of patients involved in RCTs. Rare and serious adverse events are much more likely to be reported from observation. A detailed review of adverse events was beyond the brief for this study.

Application to chronic pain

In chronic pain relief, just as in other therapeutic areas, there are often many ways to tackle a particular problem. There may be evidence of benefit for

each of the alternatives. We need a means of ranking the relative effectiveness of these interventions, so that informed decisions can be made about which should be selected, purchased and offered to patients.

Perhaps, in an ideal world, there would be large RCTs that compared the various interventions. What we have in practice is a number of small studies. The methods used to rank the relative performance of the interventions are described in later chapters. The ranking often has to be indirect, for example, how well does each intervention compare with placebo, rather than derived from direct 'head-to-head' comparisons of the treatments (*Table 1*).

TABLE 1 Ranking relative efficacy

Trial comparator	Data source	
	Individual patient data	Published group data
Comparisons with placebo	Indirect ranking	
Comparison with other 'active' interventions	Direct ranking	

Chapter 2

Finding all the relevant trials

Relavant and valid evidence is necessary to promote effective care. The RCT is the most reliable way to estimate the effect of an intervention. The principle of randomisation is simple. With randomisation, those taking part in a 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 reduces the risk of imbalance of other factors between treatment groups. Inadequate randomisation, or inadequate concealment of randomisation, leads to exaggeration of therapeutic effect.¹

Identifying all the relevant unbiased RCTs for scientifically valid reviews of evidence (systematic qualitative reviews and meta-analyses) remains a 'fundamental challenge'² whose scale is frequently underestimated.

The first obstacle faced by any reviewer trying to meet the challenge is that, in most cases, the total number of eligible RCTs is unknown. Perhaps only with new interventions can reviewers be sure that they have included all the eligible RCTs. Otherwise the total number of trials can only be identified by scanning each record in each of the available bibliographic databases, by searching manually all the journals, theses, books of proceedings, and textbooks not indexed in any bibliographic database, by searching the reference lists of all the relevant reports identified by such searches, and by obtaining all the relevant unpublished information from all the investigators who had been involved in eligible RCTs.³ In practice, however, constrained by time and by cost, reviewers have to identify the maximum possible number of eligible RCTs, and hope that it will be a representative sample of the (unknown) total population of eligible RCTs.

The failure to identify reports which could have affected the results of a systematic review or meta-analysis has been called 'retrieval bias'.⁴ Reports may not be identified because the trials are still ongoing, completed but as yet unpublished (publication bias) or are published but the methods used to identify them have failed.

The more comprehensive the searching, the more trials will be identified and the stronger

the foundations on which the conclusions will rest. Comprehensive searches, however, can be very time-consuming and costly. Reviewers must therefore decide, given their resources, which methods to use to obtain the highest possible yield.

Experience to date suggests that attempts to identify unpublished trials by undertaking surveys of large populations of researchers have such a low yield that efforts in this direction seem unjustified.⁵

Another source of unpublished data are registers of ongoing and completed trials, but these are not available in pain research.

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.

A report was regarded as eligible if 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 given elsewhere.⁶ Since that publication, the major changes are in the use of other databases in addition to Medline. Searching Embase, the Cochrane Library, CINAHL and

Psychlit is now part of our standard operating procedure (see *Figure 3*).

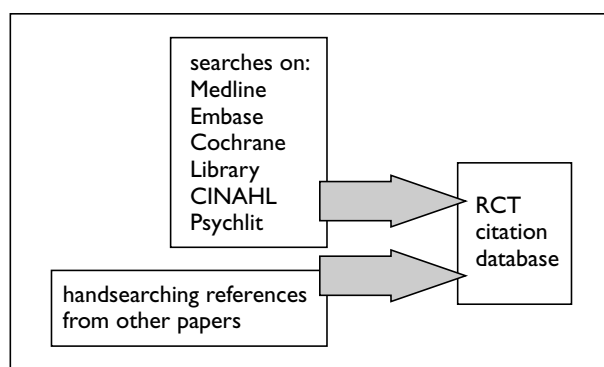


FIGURE 3 Finding the citations

Medline search for RCTs published from 1966 to date

The records identified by the optimised Medline search strategy (*Table 2*) were downloaded (Biblio-Link v. 1.1, Personal Bibliographic Software, Inc.) and transferred to a reference management program (Pro-Cite, Personal Bibliographic Software, Inc., v. 2.1). These records were then sorted in alphabetical order and each record was checked on-screen for definite eligibility, probable eligibility or ineligibility, and coded accordingly within each Pro-Cite record. Hard copies of all 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 reports of relevant RCTs. These studies, either missed by Medline indexing, or in non-indexed journals, were then added to the citation database if perusal of a hard copy confirmed that they were indeed RCTs.

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 number of RCTs in pain relief research published from 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 those identified were published in the last 15 years. This is illustrated by the trend in the

TABLE 2 Refined high-yield Medline search strategy

Step	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	RANDOMISED 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

number of RCTs reported 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 as psychological and complementary.

Conclusions

The importance of basing systematic reviews on the highest quality evidence (randomised trials) is obvious

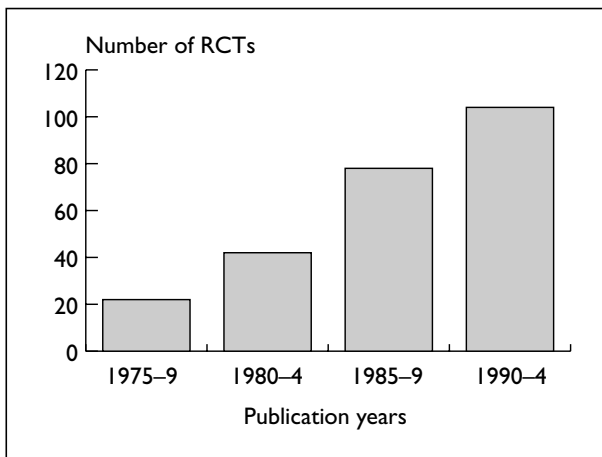


FIGURE 4 RCTs in the journal Pain from 1975-1994

from our experience in the pain field (see chapter 8) 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

TABLE 3 Breakdown of RCTs by clinical setting

Pain trials database 1950-1994					
	Acute	Chronic	Cancer	Total	%
Complementary	112	223	10	345	2
Invasive	1697	336	34	2067	14
Pharmacological	5390	4978	337	10705	75
Physical	402	501	36	939	7
Psychological	100	191	10	301	2
Total	7701	6229	427	14357	
Percentage	54	43	3		

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 (NLM), to ensure that all the RCTs found only by hand-searching are tagged.

Chapter 3

Judging the quality of trials

Once all the reports of the relevant trials have been found there is another stage in the process. This is to confirm, first, that 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. You then discover that 20 of the reports say that the intervention is terrific, and 20 conclude that it should never be used. Delving deeper you find that the 20 'negative' reports score highly on your quality standards scale, but the 20 'positive' reports score poorly. The quality scale should include measures of bias. Bias is the simplest explanation of 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 which we may want answered. In the pain world, however, there are two reasons for setting this high standard, and requiring trials to be randomised. The first is that there are, particularly for drug interventions, quite a number of RCTs. The second, we would argue, is 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.⁷ 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.⁸⁻²² 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 believe 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 the scale is to assess the likelihood of the trial design generating unbiased results and approaching the 'therapeutic truth'. This has also been described as 'scientific quality'.²³ 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 these, the judges used both previously published criteria and their own judgement. The suggestions were then combined into a single list of 49 items, as shown in *Figure 5*.

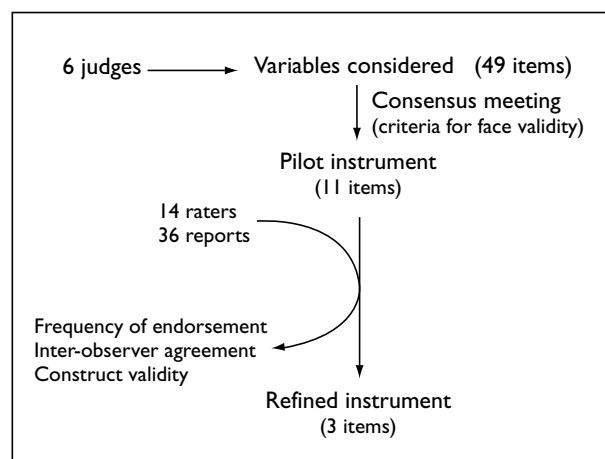


FIGURE 5 Process of developing the scale

Using a modified nominal group approach to reach consensus,²⁴ the judges assessed the face validity of each of the items, according to established criteria.²⁵ Those 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. The raters identified problems in clarity and/or application of each of the items. 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.

Final version of the scale

The final version of the scale has the three items with highest frequency of endorsement (see *Table 4*). The advice on using the scale is shown in *Table 5*.

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.⁷ This has important implications, because the cost of organising a 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 without clinical or research experience in pain relief, people should be able to score the quality of research reports consistently. Our purpose was to allow us to undertake differential analysis within our systematic reviews, based on the quality of the individual primary studies; however, the scale may have much wider use.

Chalmers suggested many years ago that the quality of clinical reports should be assessed blind.⁹ We found that such blinded assessment produced 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.^{22,26} The results of open evaluations

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.	
<i>Please read the article and try to answer the following questions (see attached instructions).</i>	
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	
<i>Award a score of 1 point for each 'yes' and 0 points for each 'no'. There are no in-between marks.</i>	
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

<p>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.</p>
<p>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.</p>
<p>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.</p>

are good enough for busy readers. The improved reliability with blind testing is more important to journal editors, for manuscript selection, and to systematic reviewers. Quality scales without clinimetric evaluation have already been used in pain studies to support the conclusions of systematic reviews.^{17,19,20}

None of the items on the scale is 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,¹⁴ 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 given consent 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.²⁷ 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 have a higher risk of selective selection.

All these methods are regarded as appropriate for the purposes of our scale, although we are aware that selective selection is still possible, even if the group allocation is concealed until after consent has been obtained. We rate the randomisation method 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 withdrawal and drop-outs occur after the allocations have been made.²⁸ This is why an adequate description of withdrawals and drop-outs is included in the scale. With that information it is possible to analyse on an intention-to-treat basis (all those randomised whether or not they were exposed to the study interventions).²⁹

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 blinding of both 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 transcutaneous electrical nerve stimulation (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%.³⁰

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.

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 neuro-physiological tests. Measurement of pain must therefore rely on recording the patient's report. The assumption is often made that because this 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 situations, 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.³¹ The patient picks the most appropriate word to describe the pain. Most research groups use four words (none, mild, moderate and severe). Scales to measure pain relief were developed later. The commonest is the five category scale (none, slight, moderate, good or lots, and complete).

For analysis, numbers are given to the verbal categories (for pain intensity, none = 0, mild = 1, moderate = 2, and severe = 3; and for pain relief, none = 0, slight = 1, moderate = 2, good or lots = 3,

and 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.^{32–34} 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 end labelled 'no relief of pain' and right end labelled 'complete relief of pain', seems to overcome this limitation. Patients mark the line at the point which corresponds to their pain. The scores are obtained by measuring the distance between the no relief end and the patient's mark, usually in millimetres. 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.

Pain relief scales are perceived as more convenient than pain intensity scales, probably because patients have the same baseline relief (zero) although they may start with a different baseline intensity (usually moderate or severe). Pain relief scale results are then easier to compare. They may also be more sensitive than pain intensity scales.^{34,35} A theoretical drawback of 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 as 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; for pain relief, 0 represents none and 10 complete relief). They are very easy and quick to use, and correlate well with conventional VASs.³⁶

Global subjective efficacy ratings, or simply global scales, are designed to measure overall treatment performance. Patients are asked questions like, ‘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?’ Its advantage is 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 like the McGill are also used. The limitation to guard against is that they usually reflect other experiences as well as or instead of pain.³⁷

Judgement by the patient rather than the carer is the ideal. Carers overestimate 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 intensity (sum of pain intensity differences (SPID)) or relief (total pain relief (TOTPAR)) measures is derived.

$$\text{SPID} = \sum_{t=1}^n \text{PID}_t \quad \text{TOTPAR} = \sum_{t=1}^n \text{PR}_t$$

Where, at the t th assessment point, ($t = 0, 1, 2, \dots, n$) P_t and PR_t 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 do not provide 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.

Standardising the summary measures

The method used to standardise TOTPAR values derived from a categorical verbal rating scale of pain relief (catPR) is shown in *Figure 6*. 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.

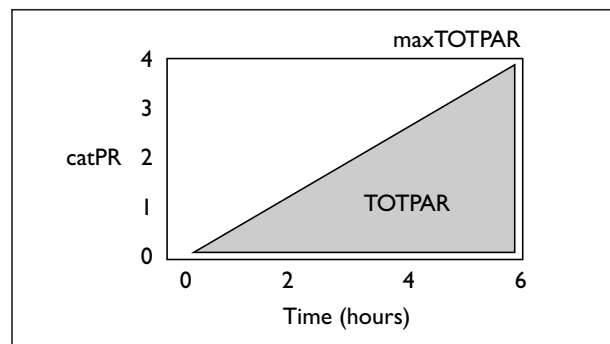


FIGURE 6 Calculating percentage of possible pain relief score

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 an understanding that the design of studies contributed directly to the validity of the result obtained. Trial designs that 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. The effect is reproducible, and some work has been done to try and assess the characteristics of the ‘placebo responder’, by sex, race 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, an RCT should prove 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 (e.g. 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 (e.g. morphine) could be used – showing the higher dose to be statistically superior to the lower dose again provides confidence that the assay is sensitive.

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

Studies of analgesics of an A versus B design are notoriously difficult to interpret. If there is a statistical difference, then that suggests sensitivity. Lack of a significant difference means nothing – there is no way in which to determine if there is an analgesic effect that is the same for A and B, or if the assay lacks the sensitivity to measure a difference that is actually present.

Equivalence

Equivalence is a more difficult problem, if only because of the large variations that occur in pain studies.^{38,39} Equivalence studies might take the form of two doses of a new analgesic compared with two doses of a standard analgesic, plus placebo to establish sensitivity. 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 complicated 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; this is often the case in chronic pain. Clearly, then, 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 is little problem with using placebos, since the failure of placebo (or any treatment) can be dealt with by prescribing additional analgesics which should work.

When there is no pain to begin with

Clearly, when 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.

Chapter 5

Estimating relative effectiveness

Minor analgesics, such as paracetamol or ibuprofen, and combination analgesics with opioids, such as codeine or dextropropoxyphene, are often used to treat chronic pain. There are few direct comparisons of one minor analgesic against another but most trials contain a placebo, which has the potential to be the universal comparator.

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.⁴⁰ 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).⁴¹ As Wall points out, these ideas stem from a misreading of Beecher's work some 40 years ago.⁴⁰

In Beecher's five acute pain studies, 139 (31%) of 452 patients given placebo had 50% or more relief of postoperative pain at two checked intervals.⁴¹ The proportion of patients with 50% or more relief from 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.⁴³

Randomised, double-blind trials are meant to eliminate (or at least minimise) both selection bias and observer bias; Evans' observation suggests that significant observer bias occurs. Wall⁴⁰ rightly questions the blindness of these trials if this result

were true, and elegantly dissects the areas where 'leakage' of blinding can occur (patient–patient; patient–doctor; patient–nurse).

Methods

Individual patient data were used from five placebo-controlled double-blind RCTs which investigated the analgesic effects of various drugs in postoperative pain^{44–48} and were performed over a 10-year period by the Pain Research Group, Oxford. All were randomised, double-blind and parallel-group trials of single doses of drugs given orally. Randomisation was from random number tables. Drugs were prepared outside the hospital in which the studies were performed. 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. 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).⁴⁹ Statistical analysis was undertaken, using Statview v. 4.02 on a Macintosh IIfx.

Results

In the five trials, 130 patients were given 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 7* for the active drugs. In the five trials, 395 patients were given active drugs. Individual

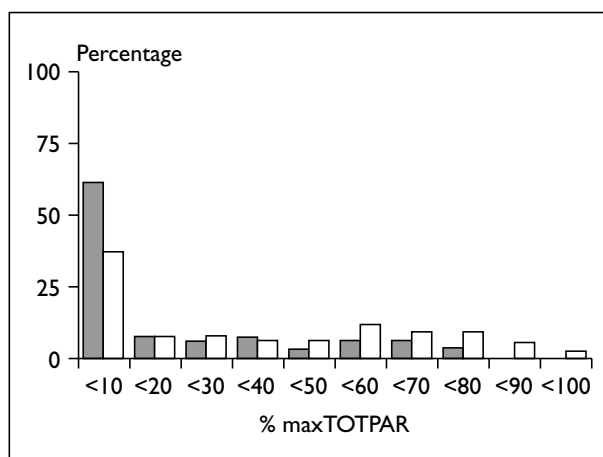


FIGURE 7 Distribution of % maxTOTPAR scores for 130 patients given placebo (■) and 395 patients given active drugs (□)

patients' scores with different active drugs varied from 0% to 97% of maximum possible pain relief.

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 8*. Mean placebo scores were related to the mean score for the active drugs in each trial such that the higher the active drug score, the higher the placebo score. A similar relationship obtained for the best and worst active drug scores from each of the five trials. On average, the mean placebo results were 54% of the mean active drug results based on a slope of 0.54; 95% confidence interval (CI) around the slope: 0.03–1.08.

The relationship between the median scores for active treatment and placebo are also shown in *Figure 7*. 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 but with a 95% CI of –0.24–+0.48, and included no relationship

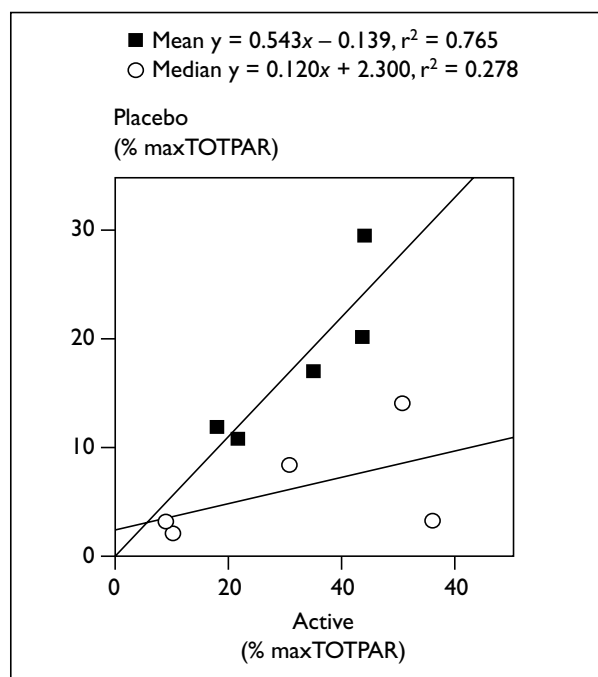


FIGURE 8 % maxTOTPAR placebo and active scores for the five trials

between placebo response and extent of the response to active analgesic.

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 in 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 greater than 50% relief in five acute pain studies.⁴² Here, using the derived dichotomous measure of 50% maximum pain relief, a range of 7–37% of patients given placebo achieved better than 50% pain relief across the five studies (see *Table 6*).

TABLE 6 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 ⁴⁴	21	11.9 (19.3)	3.1 (16.4)	2	10
Evans et al ⁴⁵	30	29.4 (29.1)	14.0 (53.0)	11	37
McQuay et al ⁴⁶	19	20.1 (29.1)	3.1 (27.3)	4	21
McQuay et al ⁴⁷	30	10.7 (17.8)	2.1 (8.3)	2	7
McQuay et al ⁴⁸	30	16.9 (21.2)	8.3 (25.0)	2	7

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 (SDs) 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 active treatment (see *Figure 7*). The predominant 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 descriptor is not valid and the use of a median is more sensible. Averaging results to describe them is a historic hangover.

In describing the placebo groups, therefore, the range of mean placebo response of 11–29% of maximum (*Table 6*) 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 which led to the observation⁴³ 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 8*) 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.^{50,51} A small number of patients may have had opportunities to

communicate with each other. Doctors who knew the trial design obtained consents from the patients, and this may also be a source of bias.⁵² 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 her demeanour, based on her experience of other patients' reactions. That would produce time-dependent changes in study results as has been observed before.⁵³

Bias may still occur but its effects are slight. That 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 non-steroidal anti-inflammatory drugs (NSAIDs) in active and placebo-controlled studies,⁵⁴ showing that the presence of 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 SD 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^{55,56} 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 number-needed-to-treat (NNT) calculations. The reality is that individual patient data are 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.⁵⁷

A proposed solution

We examined the hypothesis 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; and
- (ii) knowing that relationship allows for 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. At its simplest, 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 half 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 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. 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. Such an 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 which 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 which investigated the analgesic effects of various drugs in postoperative pain.^{44–48, 58–65} These trials were performed 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 given orally, except sublingual buprenorphine,⁵⁹ and intramuscular opioids.⁶⁵

All the studies were randomised, double-blind and parallel-group. Patients were told about the study by the nurse observer on the day before surgery. Informed consent was obtained by the doctor that evening. Randomisation was undertaken using random number tables. Drugs were prepared outside the hospital in which the studies were performed. Treatment codes were not broken until the studies were finished. 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 point. In placebo-controlled trials, all patients knew that a 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 ensured 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.⁶⁶

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 > 50% maxTOTPAR (or percentage > 50% maxTOTPAR to accommodate unequal group sizes). The dichotomous descriptor of > 50% maxTOTPAR was chosen because it is a simple clinical endpoint of pain half relieved, easily understood by professionals and patients alike.

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

Statistical analysis was performed using Statview v. 4.1 on a Macintosh IIfx. Odds ratios and their 95% CIs were calculated from standard formulae, incorporating a fixed-effects model and NNT, using the method of Cook and Sackett.⁶⁷ 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. 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 (SD, 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 > 50% maxTOTPAR using the techniques developed for the 45 actual treatments and, for these simulated treatments, to compare the calculated number with > 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 an SD of 3. These were calculated by transforming the $U[0, 1]$ values into normal values with the required mean and SD, using the Box–Mueller algorithm.⁶⁸ If any generated value of group size was less than 15 it was discarded and a new value generated which fell in 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 any value greater than 100 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 10–100. Standard techniques were then used to show that a distribution generated in this way had a theoretical mean of 37.1 and an SD 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 $U[0, 74]$, with a mean of 37.

Stage I results

The actual trials used in the analysis, the treatments used, numbers in each group, mean % maxTOTPAR and numbers of patients with > 50% maxTOTPAR are shown in *Table 7*.

TABLE 7 Studies

Study	Treatment (N)	Mean % maxTOTPAR	Actual number with > 50% maxTOTPAR	Calculated number with > 50% maxTOTPAR
Evans et al, 1982 ⁴⁵	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 ⁶⁰	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, 1987a ⁴⁷	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 ⁴⁴	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 ⁴⁸	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 ⁵⁹	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 ⁵⁸	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
	placebo (30)	16.9	2	3
McQuay et al, 1985 ⁴⁶	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 ⁶³	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
	placebo (30)	16.9	2	3
McQuay et al, 1993 ⁶⁴	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 7 contd Studies

Study	Treatment (N)	Mean % maxTOTPAR	Actual number with > 50% maxTOTPAR	Calculated number with > 50% maxTOTPAR
McQuay <i>et al</i> , 1989 ⁶²	ibuprofen, 400 mg (23)	44.8	10	11
	ibuprofen, 400 mg + codeine, 20 mg (24)	57.7	15	16
McQuay <i>et al</i> , 1987b ⁶¹	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 <i>et al</i> , unpublished ⁶⁵	pethidine, 100 mg (21)	16.1	1	2
	meptazinol, 100 mg (20)	21.8	1	3
	morphine, 15 mg (22)	24.4	4	4

The calculated number of patients in each treatment group with > 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 more than half relief. Numerical values were rounded up or down to the nearest integer.

Actual mean and proportion with > 50% maxTOTPAR

The relationship between mean % maxTOTPAR and proportion with > 50% maxTOTPAR is shown in *Figure 9*. The equation to the regression line was:

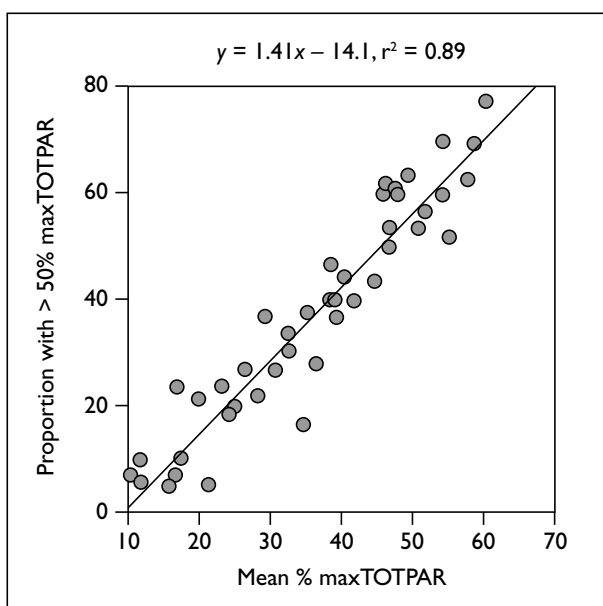


FIGURE 9 Relationship between mean % maxTOTPAR and proportion with > 50% maxTOTPAR

$$\begin{aligned} \text{Percentage of patients} &= 1.41 \% \text{ maxTOTPAR} - 14.1 \\ > 50\% \text{ maxTOTPAR} & \quad (r^2 = 0.89) \end{aligned}$$

Calculated number of patients with > 50% maxTOTPAR

The actual and calculated numbers of patients in each group with > 50% maxTOTPAR are shown in *Table 7*. The equation to the regression line was:

$$\begin{aligned} \text{Calculated number of patients} &= 0.93 \text{ actual} + 0.93 \\ > 50\% \text{ maxTOTPAR} & \quad (r^2 = 0.88) \end{aligned}$$

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 4 patients. The two most aberrant results occurred in the same trial.⁶¹

Simulated actual distribution – mean and proportion with > 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 of patients > 50% maxTOTPAR which was very similar to that obtained for the actual data from 45 treatments.

$$\begin{aligned} \text{Percentage of patients} &= 1.34 \text{ mean \% maxTOTPAR} \\ > 50\% \text{ maxTOTPAR} & \quad - 14.1 \quad (r^2 = 0.79) \end{aligned}$$

Simulated actual distribution – calculated numbers > 50% maxTOTPAR

The equation above 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.

$$\begin{aligned} \text{Calculated number of patients} > 50\% \text{ maxTOTPAR} &= 0.82 \text{ actual} + 1.92 \\ &\quad (r^2 = 0.83) \end{aligned}$$

From *Table 8*, it can be seen that, using the underlying distribution, the difference between calculated and actual number of patients with > 50% maxTOTPAR was 0–2 for 90% of the simulated studies and, for 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 of patients with > 50% maxTOTPAR was repeated with two further distributions. The results obtained for the ‘normal’ and ‘uniform’ distributions (see *Table 8*) 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 9*). NNT values

TABLE 8 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 9 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)

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 NNTs for analgesic effectiveness.⁶⁷ At least two identical treatments from different trials were required. Of the 130 patients who received placebo, 21 actually had > 50% maxTOTPAR and 15 were calculated to have > 50% maxTOTPAR.

Verification from independent data

Stage 2 methods

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

Study protocols for post-surgical 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. Patient ages ranged from 18 to 70 years. Patients had to be co-operative, 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 who needed sedatives during the observation period, and those with known contraindications or medical conditions that 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); acetaminophen, 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 post-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, as well as 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, and % maxTOTPAR was then calculated for each patient.⁴⁹ Mean TOTPAR was calculated for all patients in each treatment arm, and the number of patients on each treatment achieving > 50% maxTOTPAR was noted.

The mean TOTPAR value was then used to calculate the theoretical number of patients with > 50% maxTOTPAR using a relationship established in clinical trials of analgesics at Oxford with 1283 patients and 45 treatments (percentage of patients with > 50% maxTOTPAR = 1.41 x mean % maxTOTPAR - 14.1). Actual and calculated numbers were then compared by unweighted linear regression analysis using Excel v. 5 and Cricket Graph v. 1.5 on a Power Macintosh 7100/80.

Stage 2 results

Individual patient information was available from over 3400 patients in 85 different treatment arms of nine studies involving dental surgery (mostly third molar extraction) and nine involving general postoperative pain (including gynaecological procedures). Studies included 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 10*.

The relationship between actual and calculated numbers of patients with > 50% maxTOTPAR in each treatment arm is shown in *Figure 11*, and the equation to the regression line for this is compared in *Table 10* with 45 treatments from trials in Oxford, 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

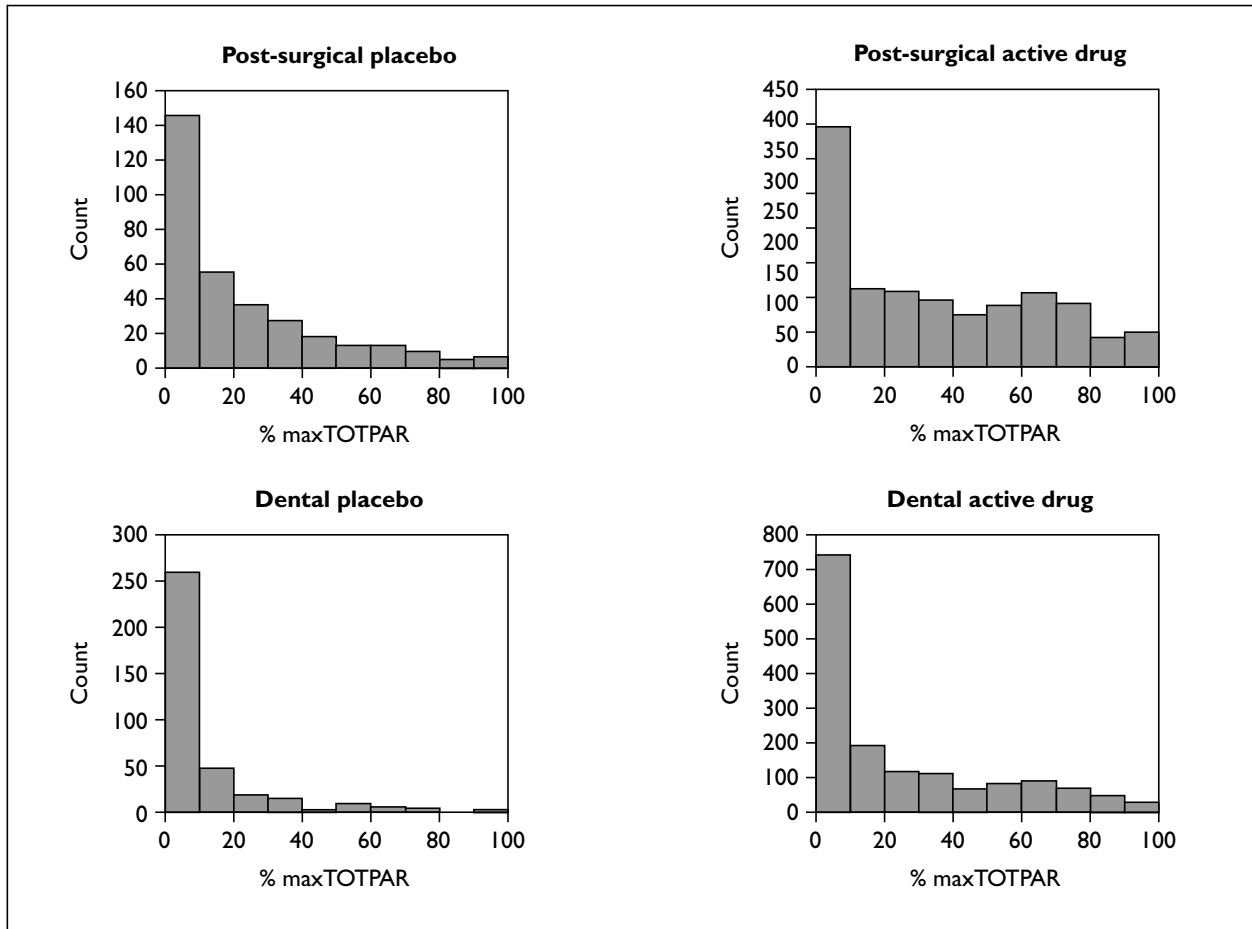


FIGURE 10 Distribution of % maxTOTPAR for active treatments and placebo in dental and post-surgical pain

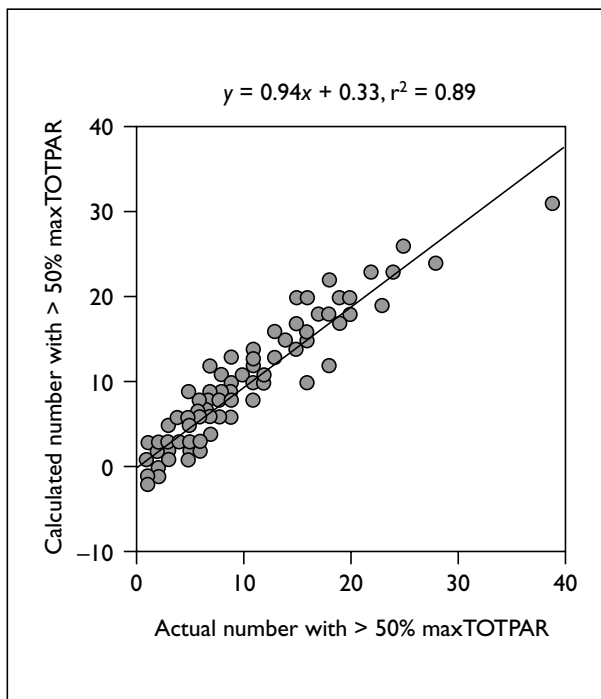


FIGURE 11 Relationship between actual and calculated number with at least 50% maxTOTPAR for 3400 patients with 85 treatments

three (Table 11). These proportions are comparable to those obtained previously for actual and simulated treatments (see Table 8). Summing the positive and negative differences between actual and calculated numbers of patients with > 50% maxTOTPAR gave an average difference of 0.30 patients per treatment arm.

Comparison of actual:calculated (irrespective of sign) numbers of patients with > 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 11.⁵⁷ 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 treatments⁵⁷ produced a new relationship to be used in future conversions.

$$\begin{aligned} \text{Proportion of patients with } > 50\% \text{ maxTOTPAR} &= 1.33 \times \text{mean } \% \text{ maxTOTPAR} - 11.5 \\ & \quad (r^2 = 0.89) \end{aligned}$$

TABLE 10 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 11 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

Use of pain intensity and VAS

Stage 3 methods

Data for the study were taken from individual patient data from 13 RCTs (1283 patients with 45 treatments, Oxford data⁵⁷) and 18 RCTs (3453 patients with 87 treatments, RWJ data) described in Stages 1 and 2.

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 an 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.⁴⁹ 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 > 50% maxTOTPAR was noted.

The relationship between the mean % maxSPID, % maxVAS-SPID and % maxVAS-TOTPAR and the actual number of patients with > 50% maxTOTPAR was examined by linear regression analysis. Using the equation to the regression line, the calculated number of patients with > 50% maxTOTPAR was then compared with the actual number using unweighted linear regression analysis. The calculations were performed using Excel v. 5, StatView v. 4.51 and Cricket Graph v. 1.5 on a Power Macintosh 7100/80.

Stage 3 results

Individual patient scores for categorical pain intensity and visual analogue pain intensity and pain relief scales 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 12*). Linear regression analysis performed for the Oxford

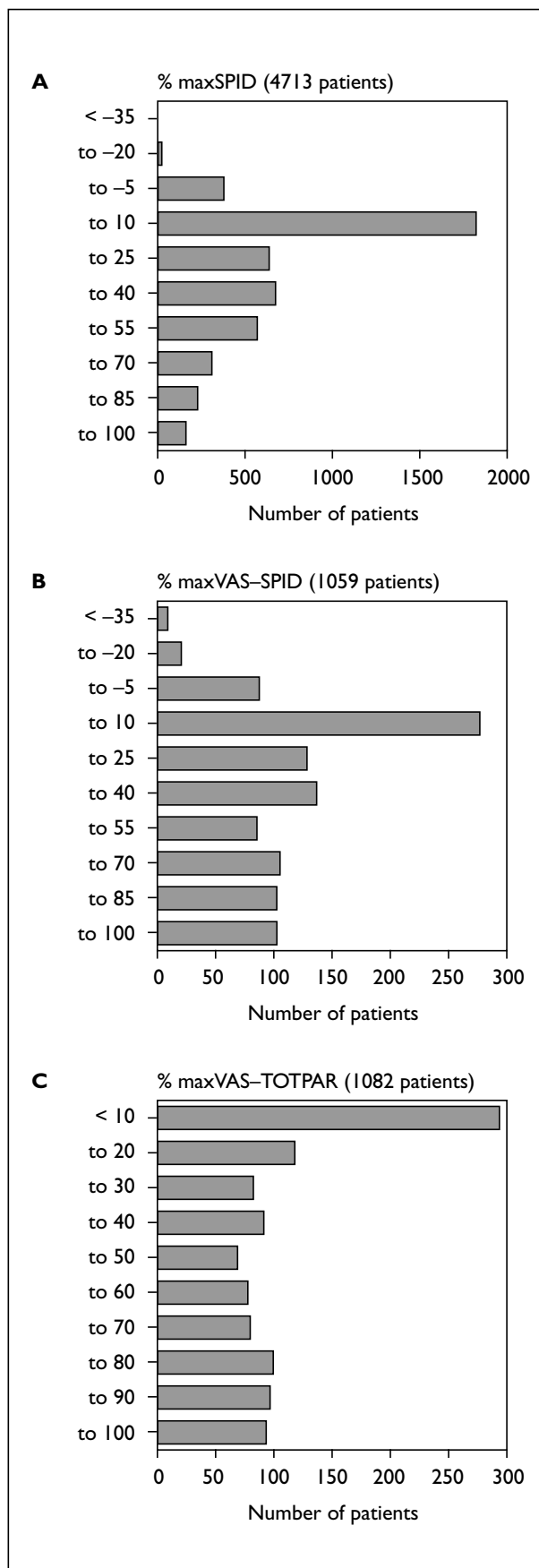


FIGURE 12 Distribution of individual patient scores for categorical pain intensity and VAS pain intensity and pain relief scores

and RWJ data sets separately showed similar relationships, so the data sets were combined for all 132 treatments.

For all 132 treatments, the regression line was:

$$\text{Percentage with } > 50\% \text{ maxTOTPAR} = 1.36 \text{ mean } \% \text{ maxSPID} - 2.3 \quad (r^2 = 0.85)$$

There was good agreement between the actual numbers of patients with $> 50\%$ maxTOTPAR in each treatment arm and the numbers calculated using the relationship derived with % maxSPID (*Figure 13*):

$$\text{Calculated number with } > 50\% \text{ maxTOTPAR} = 0.86 \text{ actual} + 1.37 \quad (r^2 = 0.86)$$

For 92% of treatments, the actual and the calculated numbers with $> 50\%$ maxTOTPAR were within four patients per treatment. Agreement (actual:calculated) was normally distributed around zero (*Figure 14*). Summing the positive and negative differences between actual and calculated numbers of patients with $> 50\%$ 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 (see *Figure 13B*). The regression line between percentage with $> 50\%$ maxTOTPAR and mean % maxVAS-SPID was given by:

$$\text{Percentage with } > 50\% \text{ maxTOTPAR} = 1.18 \text{ mean } \% \text{ maxVAS-SPID} - 2.2 \quad (r^2 = 0.87).$$

There was good agreement between the actual number of patients with $> 50\%$ maxTOTPAR in each treatment arm and the number calculated, using the relationship derived from % maxVAS-SPID (see *Figure 13B*):

$$\text{Calculated number with } > 50\% \text{ maxTOTPAR} = 0.90 \text{ actual} + 1.19 \quad (r^2 = 0.79).$$

For 95% of treatments, the actual and the calculated numbers with $> 50\%$ maxTOTPAR were within four patients per treatment. Summing the positive and negative differences between actual and calculated numbers of patients with $> 50\%$ maxTOTPAR gave an average difference of -0.23 patients per treatment arm.

VAS for pain relief

Data were available from 40 treatments with 1082 patients. Individual patient distribution of % maxVAS–TOTPAR was asymmetric (see *Figure 12C*). The regression line between percentage with > 50% maxTOTPAR and mean % maxVAS–TOTPAR was given by:

$$\begin{aligned} \text{Percentage with } > 50\% \text{ maxTOTPAR} &= 1.15 \text{ mean } \% \text{ maxVAS–TOTPAR} \\ &- 8.51 \quad (r^2 = 0.81). \end{aligned}$$

There was good agreement between the actual number of patients with > 50% maxTOTPAR in each treatment arm and the number calculated, derived from % maxVAS–TOTPAR (see *Figure 12C*):

$$\begin{aligned} \text{Calculated number with } > 50\% \text{ maxTOTPAR} &= 0.89 \text{ actual} + 1.15 \\ &\quad (r^2 = 0.81). \end{aligned}$$

For 95% of treatments, the actual and the calculated numbers with > 50% maxTOTPAR were within four patients per treatment. Summing the positive and negative differences between actual and calculated numbers of patients with > 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, using independent data sets. Regressing % > 50% maxTOTPAR against mean % maxSPID independently for Oxford and RWJ data sets produced very similar results (see *Table 12*). Using combined regression analysis, there was excellent agreement between actual and calculated numbers of patients with > 50% maxTOTPAR for each treatment (see *Figure 13A*), and the sum of the difference over all 132 treatments was –0.03 patients per treatment, with the differences distributed normally around zero (see *Figure 14*). 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 > 50% maxTOTPAR and mean % maxVAS–SPID and mean % maxVAS–TOTPAR. Despite this, the agreement between actual and calculated numbers with > 50% maxTOTPAR was good (see *Figure 13B, C* and *Table 13*), so that over the 40 treatments the sum of actual – calculated was less than a quarter of a patient per treatment arm using either measure.

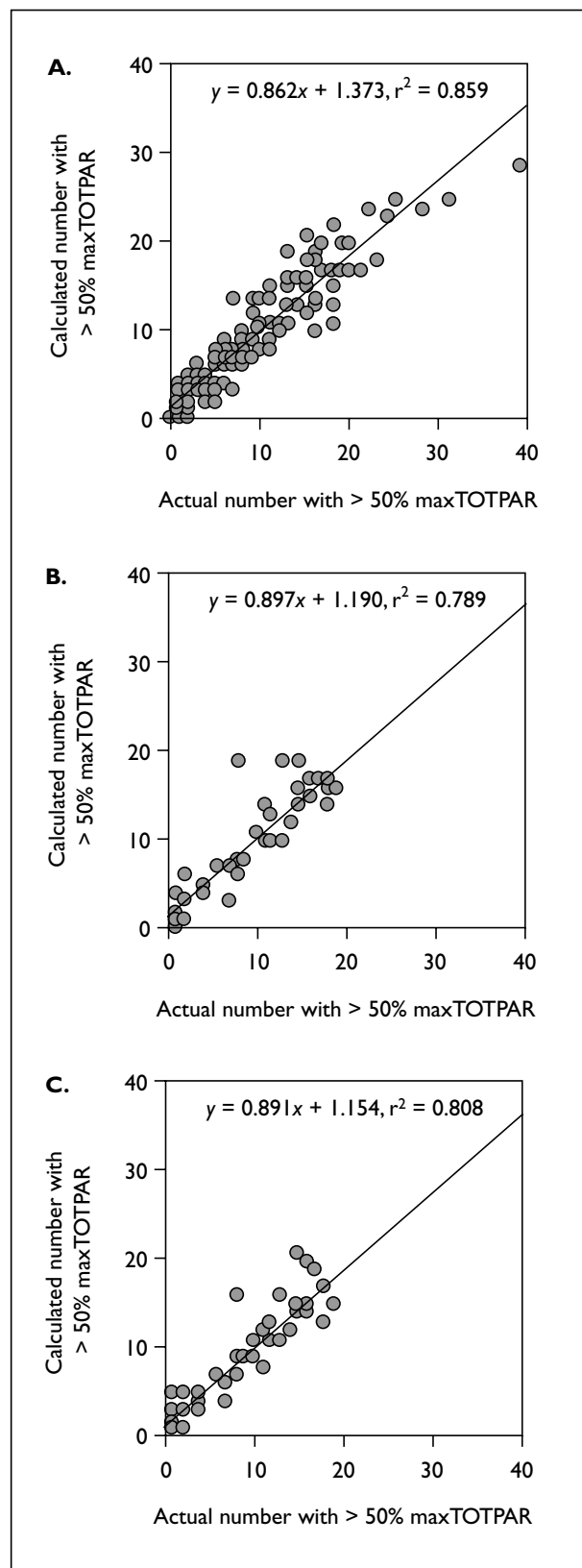


FIGURE 13 Correlation of actual and calculated numbers of patients with > 50% maxTOTPAR in each treatment arm for calculations using categorical pain intensity and VAS pain intensity and relief scores. (A) From SPID conversion in 132 treatments; (B) from VAS–SPID conversion in 40 treatments; (C) from VAS–TOTPAR conversion in 40 treatments

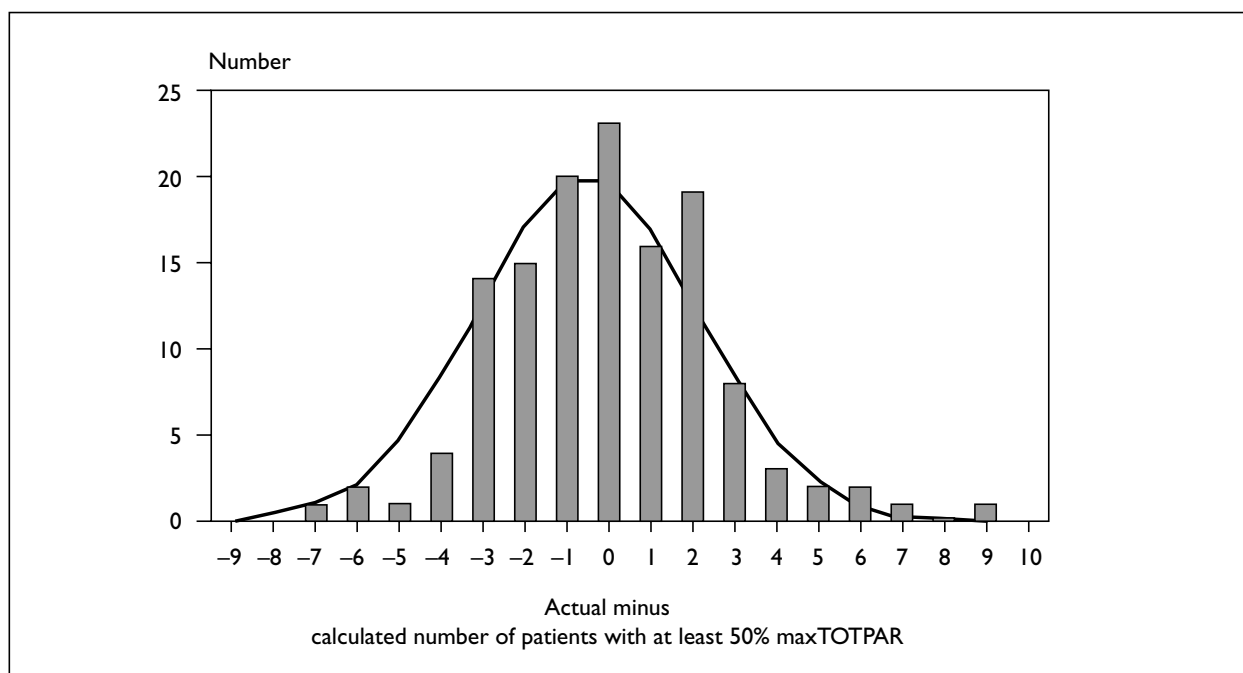


FIGURE 14 Normal distribution of actual – calculated number of patients with > 50% maxTOTPAR in each treatment using SPID data

TABLE 12 Summary report on SPID calculations on 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

TABLE 13 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

Comparison of actual minus calculated (irrespective of sign) percentage of patients with > 50% maxTOTPAR as cumulative percentages for the 45 Oxford treatments using TOTPAR,⁵⁷ and 85 RWJ treatments using TOTPAR.⁶⁹ 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 the Oxford and RWJ data sets.

Although no independent verification was possible for VAS, the similarity of the results (see *Table 13* and *Figure 13*) to those independently verified for TOTPAR^{57, 69} and SPID supports the approach of using mean data from previously published reports to derive dichotomous data for meta-analysis.⁵⁷

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.⁶⁶ 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 numbers of patients with > 50% maxTOTPAR – was essentially the same as that obtained originally using the relationship between the actual data and that from a 10,000 treatment arm simulation. Verification was also possible for SPID but not for VAS, although 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 > 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 they become available to further validate these relationships (*Table 14*), 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.

Relative efficacy of minor analgesics

The relative efficacy of minor analgesics has been explored in order to generate a developing ladder of effectiveness. Data from a number of sources have been used, including a single-patient meta-analysis of tramadol registration trials and meta-analysis of published data for paracetamol, paracetamol plus codeine, paracetamol plus dextropropoxyphene, ibuprofen and dextropropoxyphene. Their relative efficacies are shown in *Figure 15*.

It needs to be emphasised that these data are from single-dose comparisons of oral analgesics in patients with moderate or severe acute post-operative pain. The relative effectiveness of the same drugs given in multiple dosing situations, such as in chronic pain, may differ. For instance, combinations of paracetamol with opiates may be relatively more effective in chronic dosing.

Adverse effects may also play a significant role, with effectiveness, in chronic dosing situations. NSAIDs may not be indicated in elderly people because of an association with gastrointestinal bleeding – especially in patients with a history of peptic ulcer, gastrointestinal bleeding or heart disease.⁷⁰ Cost may also be an issue.

Nevertheless, information on relative effectiveness is an important part of the prescribing decision, and this information is now available.

TABLE 14 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% pain relief	Formula
Categorical pain relief		$1.33 \times \text{mean \% maxTOTPAR} - 11.5$
Categorical pain intensity		$1.36 \text{ mean \% maxSPID} - 2.3$
VAS pain relief		$1.15 \text{ mean \% maxVAS-TOTPAR} - 8.51$
VAS pain intensity		$1.18 \text{ mean \% maxVAS-SPID} - 2.2$

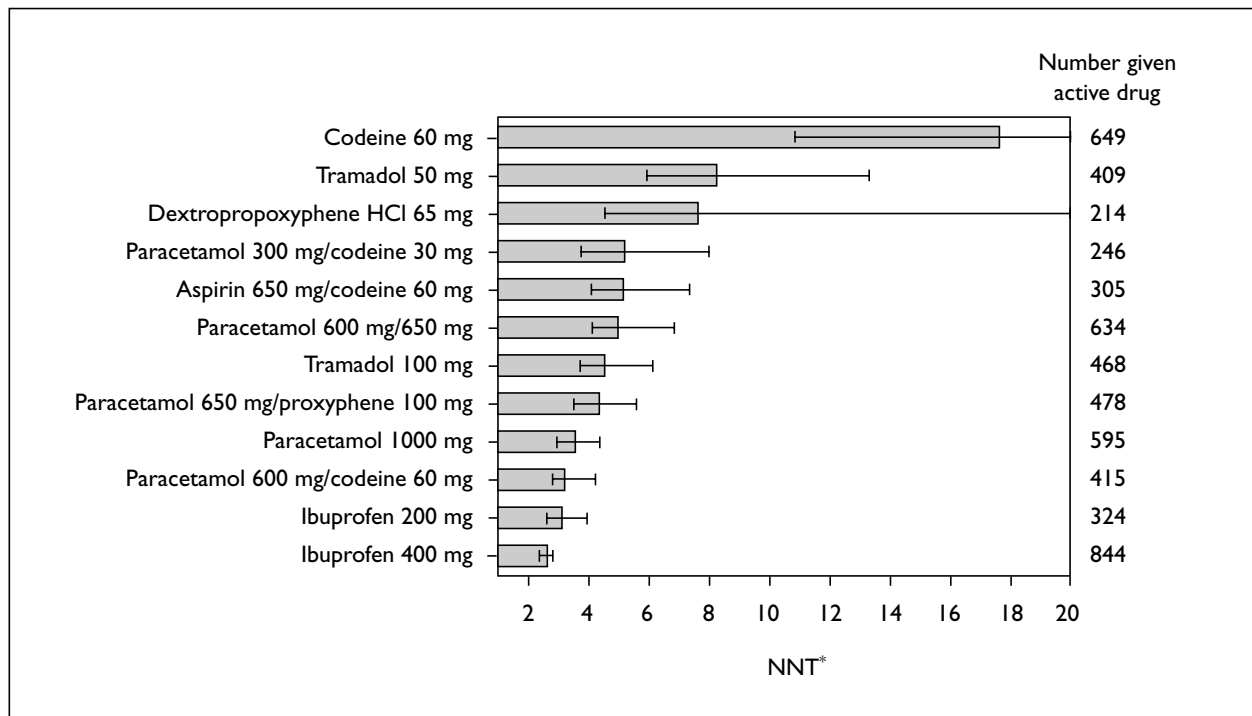


FIGURE 15 Relative efficacy of minor analgesics (*NNT for one patient with moderate or severe pain intensity to achieve >50% pain relief compared with placebo)

Chapter 6

Combining data and interpreting the 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 made 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. They use explicit methods and quality standards to reduce bias. Their results are the closest we are likely to get to the truth in the current state of knowledge.

The questions a systematic review should answer are as follows.

- How well does an intervention work (compared with placebo, no treatment or other interventions in current use) – or can I forget about it?
- 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, and 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 systematic review and particularly meta-analysis – often some sort of statistical output – 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 required. We think that common currency is 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.⁷¹

- 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 Vander Stichele *et al*⁷² for an excellent discussion of eligibility criteria for trials of head lice infection), reasons should be sought. Reviews may use criteria that exclude information important to individual clinicians, or may include studies with inadequate trial design. The 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 is appropriate. The reviewer may have used all the information that is available and any problems are due to the original trials, but it is 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?

For a systematic review of therapeutic efficacy the gold standard is that eligible studies should be RCTs. If trials are not randomised, estimates of treatment effect may be exaggerated by up to 40%.³⁰ In a systematic review of TENS in postoperative pain, 17 reports on 786 patients could be regarded unequivocally as RCTs in acute

postoperative pain. Of these, 15 RCTs demonstrated no benefit of TENS over placebo. A total of 19 reports had pain outcomes but were not RCTs; and, in 17 of these, TENS was considered by their authors to have had a positive analgesic effect.⁷³ When appropriate, and particularly with subjective outcomes, the gold standard for an efficacy systematic review is those studies which are both randomised and double-blind. The therapeutic effect may be exaggerated by up to 20% in trials with deficient blinding.³⁰

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 the trials used different clinical outcomes or followed the patients for different lengths of time. Combining continuous rather than dichotomous data may be difficult. Even if trials measure and present dichotomous data, if they are otherwise of poor quality⁷ 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 the 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 this simple vote-counting is that it may be misleading. It ignores the sample sizes, the magnitude of the effects in the constituent 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, statistical testing (odds ratio or relative risk), and a clinical significance measure such as NNT.

L’Abbé plots⁷⁴

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 a review does not show the data in this way, they can be extracted from information on individual trials presented in the review tables. *Figure 16* contains data extracted from three different systematic reviews of treatments for painful diabetic neuropathy.^{75–77} 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 control event rate (CER).

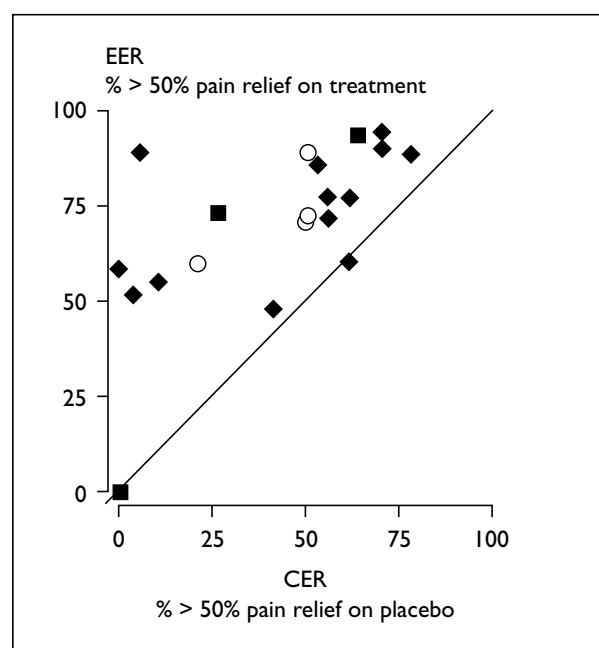


FIGURE 16 L’Abbé plot of EER (% > 50% relief on treatment) against CER (% > 50% relief on placebo) for RCTs of anti-convulsants (■), antidepressants (◆) and topical capsaicin (○) in diabetic neuropathy^{75–77}

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. All three interventions shown in *Figure 16* 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).

Visual inspection gives a quick and easy indication of the level of agreement among trials. Heterogeneity is often assumed to be due to variation in the EER, the effect of the intervention. Variation in the CER can also be a source of heterogeneity, as shown in *Figure 16* and, in this case, the controls were all matched with placebo in a relatively homogeneous chronic condition with treatment

over a period ranging from several weeks to several months.

L'Abbé plots are not yet widely used. They do have several benefits. The simple visual presentation is easy to assimilate. They make us think about the reasons why there can be such wide variation in (especially) placebo responses, and about other factors in the overall package of care that can contribute to effectiveness. They explain the need for placebo controls if ethical issues about future trials arise. They invite scepticism about overly good or bad results for an intervention in a single trial where the major influence may be how good or bad was the response 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 vomiting⁷⁸ and, in six trials of prophylactic natural surfactant for preterm infants, the CER for bronchopulmonary dysplasia was 24–69%.⁷⁹ Such variation would not always be expected, such as in the use of antimicrobials.

H. pylori eradication rates with short-term use of ulcer healing drugs were 0–17% in 11 RCTs (with 10/11 being below 10%).⁸⁰

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 variation 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 drug.

An individual patient can have no pain relief or 100% pain relief. Random selection of patients can therefore produce groups with low placebo response rate or a high placebo response rate, or somewhere in between. On-going mathematical modelling based on individual patient data shows that, while group sizes of up to 50 patients are likely to show a statistical difference for 80–90% 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 review.

Examples of the way in which 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, presented in *Figure 17*, is of trials in diabetic neuropathy where the proportion of patients given placebo is plotted against the number given placebo.

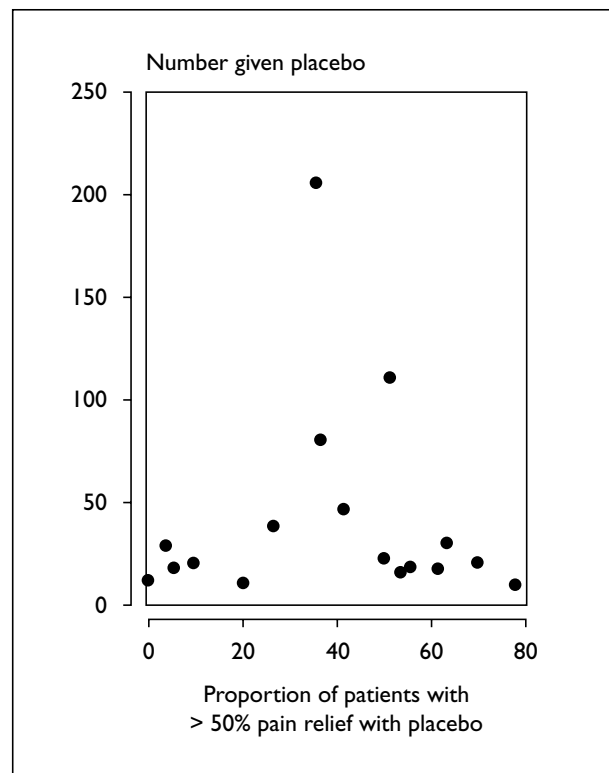


FIGURE 17 Relationship between placebo response and trial size for pharmacological interventions in diabetic neuropathy

A similar pattern of an inverted 'V' is also seen in topical NSAID trials, and indicates that almost all of the variability in placebo response occurs in trials of small size. In a study of rheumatoid arthritis, Gøtzsche⁸¹ 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 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, 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

expecting less spread of the CER may be more acceptable.⁸² In certain circumstances, for instance, in prophylaxis for nausea and vomiting, particular CER spreads may be determinants of trial validity.⁸³

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 relative risk are the accepted statistical tests to show that an intervention works significantly better than the comparator. As systematic reviews are used more often 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 18*. 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, which means 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.⁸²

Where CERs are high (certainly when they are above 50%), odds ratios should be interpreted with caution.

Relative risk

The fact that it is the odds ratio rather than relative risk reduction, which is used as the test of statistical significance for systematic reviews, seems to be due to custom and practice rather than any inherent intellectual advantage.⁸² Relative risk may be better than an odds ratio because it is more robust in situations where CER is high.⁸⁴ With event rates above 10%, relative risk produces more conservative figures.⁸⁵

In the following chapters, both odds ratios and relative risks are used, reflecting a degree of uncertainty and disagreement amongst statisticians and reviewers. In all cases, the actual numbers are given so that, when the dust has settled, calculations can be re-done 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 in standard software.⁸⁶ Unfortunately, all of these tests lack power, so while a positive test for heterogeneity suggests that mixed fruits are being compared, a negative test does not provide complete reassurance of no heterogeneity.

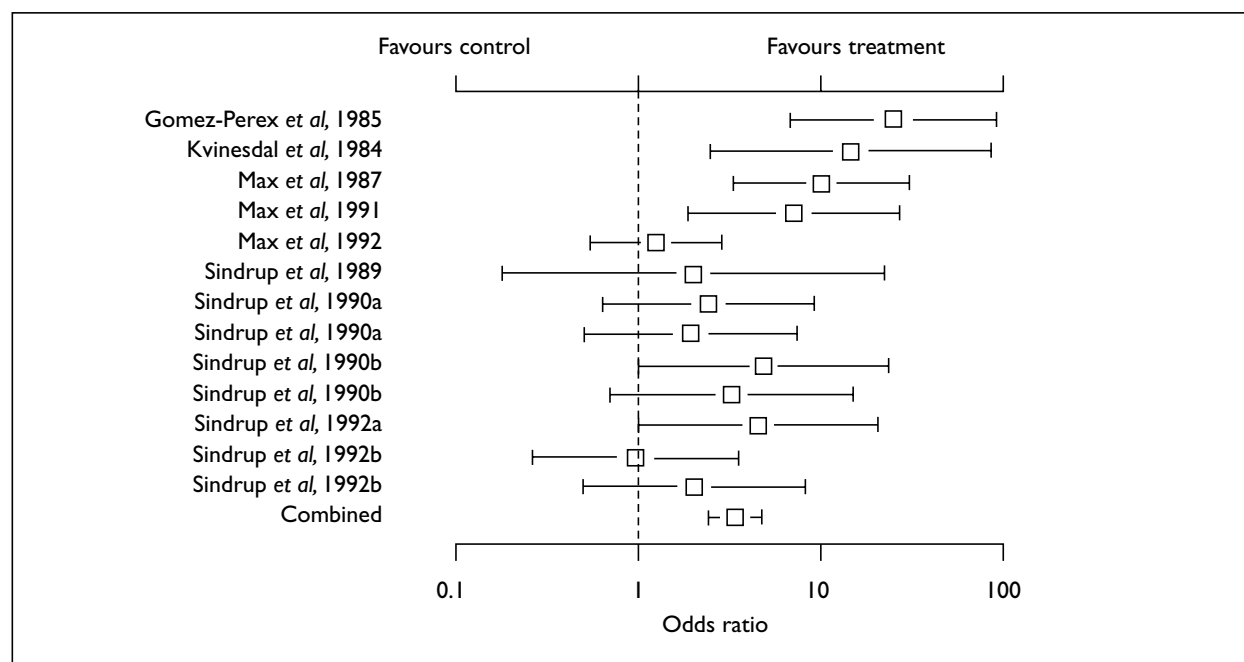


FIGURE 18 Odds ratios for antidepressants in diabetic neuropathy

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

How well does the 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 that intervention works – the size of the effect or its clinical significance.

Effect size

The classical method of estimating size of effect was to use the standardised mean difference.⁸⁷ 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 units of SD and is therefore scale-free.

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

Number-needed-to-treat

The concept of NNT 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

$$\text{NNT} = \frac{1}{(\text{IMP}_{\text{act}}/\text{TOT}_{\text{act}}) - (\text{IMP}_{\text{con}}/\text{TOT}_{\text{con}})}$$

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

Treatment-specific

NNT is treatment-specific. It describes the difference between active treatment and control. The

threshold used to calculate NNT can vary but the NNT is likely to be relatively unchanged because a varying threshold changes the results for both active and control treatment.

In the example below (*Figure 19*), from an individual patient data meta-analysis of postoperative pain relief, NNTs compared with placebo were calculated for paracetamol, 650 mg, plus dextropropoxyphene, 100 mg, for between 20% and 80% relief of pain. With placebo, the proportion of patients achieving a particular level of pain relief fell quickly as the 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.

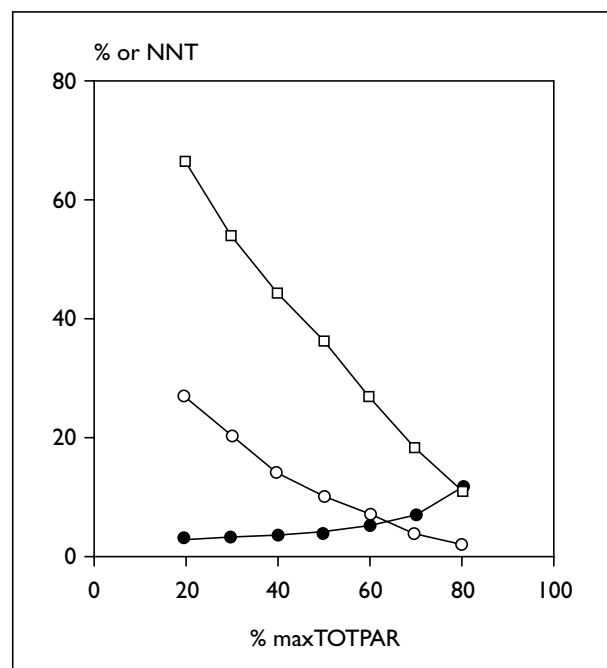


FIGURE 19 Effect of different thresholds of pain relief on NNT (□-□ Paracetamol plus dextropropoxyphene, ○-○ Placebo, ●-● NNT)

An NNT value 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 drug 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 NNT values of 2 or 3 often indicate an effective intervention. For unwanted effects, the 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).^{72, 88}

Confidence intervals

The CIs for an NNT are an indication that, in 19 cases out of 20, the ‘true’ value will be in the specified range. If the odds ratio is not statistically significant then the NNT is infinite. 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.⁵⁷ Because of the way in which it is calculated, the NNT will also be sensitive to trials with high CERs. As the CER rises, the potential for treatment-specific improvement decreases: higher (and apparently less effective) NNTs are the result. So, as with any summary measure from a quantitative systematic review, NNT needs to be treated with caution and 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 15*. This was published for prophylactic interventions⁸⁴ when the odds ratio and CER are known.

Formula for **prophylaxis**:

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

where OR = odds ratio

Formula for **treatment**:

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

where OR = odds ratio.

Choose the column in *Table 15* which is closest to the published odds ratio (**prophylaxis** left side, **treatment** right side) and the row which is closest to the event rate expected, then read off the corresponding NNT. This table can also be used to see how different values for event rate or EER for an individual patient affect the NNT at a given odds ratio.

A caveat must be added here that odds ratios should be interpreted with caution when events occur commonly, as in treatments, and odds ratios may overestimate the benefits

TABLE 15 Calculating NNTs from odds ratios

		Odds ratios																	
		Preventive									Treatment								
		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
CER	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

of an effect when event rates are above 50%. They are likely to be superseded by relative risk because this is more robust in situations where event rates are high.^{82, 89}

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 patient numbers. New medicines may be launched after trials on 1500 patients,⁹⁰ missing these 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, 3/1500.⁹¹

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

Number-needed-to-harm

For the minor adverse effects reported in RCTs, the NNH may be calculated in the same way as 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 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 evidence

The gold standard for evidence of harm, as for efficacy, is the RCT. The problem is that, in the relatively small numbers 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, simply because it is reported in a case report rather than 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.

The thrust of this report is to establish the effectiveness or otherwise of a range of interventions and, if effective, to use the NNT as a benchmark of the effectiveness of a particular intervention.

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 a particular intervention for an individual patient.

Chapter 7

Existing systematic reviews

As part of the evidence-gathering exercise we set out to find all previous systematic reviews of analgesic interventions.

This was in two stages. In the first, we sought all systematic reviews published before 1993/94. We found 80 and two judges assessed their quality using Oxman and Guyatt's index.²³ 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 on 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 of a particular topic, the results did not always agree. A full account of this first stage has been published.⁹²

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.^{93–95} 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 (Silver Platter Medline v. 3.0, 3.1 and 3.11) search was undertaken from 1966 to October 1993. This Medline strategy had been developed to identify the maximum possible number of randomised, double-blind studies or meta-analyses in pain research, and contained text words, 'wild cards' and MeSH terms.³ Forty journals were searched by hand. The register of systematic reviews at the UK Cochrane Centre was searched for eligible studies, and lead authors of abstracts were asked for full manuscripts. Reference lists were scanned for citations of other systematic reviews.

Methodological evaluation

Each study was evaluated twice, using Oxman and Guyatt's index,^{23,96} with the name of the journal, the authors, the date of publication and the source of financial support of the reports 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 both on RCTs and on systematic reviews which used other study designs. Prior hypotheses were that poor quality reviews and those designed other than as RCTs would be more likely to produce positive conclusions.

Results

Stage I: Quality assessment of reviews to 1993/94

Of the 84 reports found, 70 were included in the quality assessment. The exclusions have been specified elsewhere.⁹² The earliest report

found was from 1980, and over two-thirds appeared after 1990. Reviews considered between two and 196 primary studies (median, 28). A total of 60 reviews reached positive conclusions, seven negative, 12 uncertain and one did not reach any conclusion. All were based on published data only (no individual patient data analysis), without validity checks with the study investigators. The reviews are summarised in *Tables 16* and *17*. A separate reference list of the reviews, by first author, follows the main numerical reference list at the end of this document (see page 130).

Overall quality scores

The median agreed overall score for the systematic reviews was 4 (range, 1–7). High

TABLE 16 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 17 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)		

quality systematic reviews were significantly less likely to produce positive results (*Table 18*; chi-square = 18.2, $p = 0.006$).

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-square = 5.07; $p = 0.024$). All conclusions from systematic reviews of psychological inter-

ventions 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 score possible.

Interventions evaluated by multiple systematic reviews

There was more than one systematic review for six interventions (*Table 19*). 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.^{17,97} Three reviews confirmed that the risk of gastrointestinal complications was increased by NSAIDs.⁹⁸⁻¹⁰⁰

Most systematic reviews of manipulation for chronic back pain concluded it was useful,¹⁰¹⁻¹⁰³ as did reviews of second-line drugs for rheumatoid arthritis,¹⁰⁴⁻¹⁰⁶ but for both interventions one review questioned the validity of the findings because of the high risk of bias in the primary studies.^{20,107}

Systematic reviews produced conclusions in opposite directions for lasers in musculoskeletal pain^{108,109} and for interventions to prevent post-herpetic neuralgia.¹¹⁰⁻¹¹³ Both systematic reviews evaluating laser treatment were given the same quality score.

TABLE 18 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

TABLE 19 Multiple systematic reviews on a particular intervention: quality and conclusions

Intervention	First authors	Quality score	Conclusion
Acupuncture	Ter Riet ¹⁷	6	Uncertain
	Patel ⁹⁷	5	Uncertain
Gastrointestinal effects of NSAIDs	Chalmers ⁹⁸	7	Positive
	Gabriel ⁹⁹	6	Positive
	Bollini ¹⁰⁰	5	Positive
Manipulation	Koes ²⁰	6	Uncertain
	Shekelle ¹⁰¹	6	Positive
	Ottenbacher ¹⁰²	4	Positive
	Anderson ¹⁰³	3	Positive
Second-line drugs for rheumatoid arthritis	Gøtzsche ¹⁰⁷	7	Uncertain
	Felson (1990) ¹⁰⁵	6	Positive
	Capell ¹⁰⁴	4	Positive
	Felson (1992) ¹⁰⁶	4	Positive
Prevention of postherpetic neuralgia	Schmader ¹¹⁰	7	Uncertain
	Lycka ¹¹¹	5	Positive
	Crooks ¹¹³	3	Positive
	Naldi ¹¹²	3	Negative
Laser for musculoskeletal pain	Gam ¹⁰⁹	6	Negative
	Beckerman ¹⁰⁸	6	Positive

Comments

The use of systematic reviews to assess analgesic interventions is increasing, but most of the reviews we found had methodological flaws which may threaten their conclusions. Only eight of the 80 satisfied all the Oxman and Guyatt criteria 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.

Chapter 8

Transcutaneous electrical nerve stimulation

TENS was originally developed as a way of controlling pain through the ‘gate’ theory.¹¹⁴ According to the theory, selective stimulation of certain nerve fibres could block, or ‘close the gate’ on signals carrying pain impulses to the brain.

TENS is widely used. A survey of 50 Canadian hospitals with 200 or more beds¹¹⁵ indicates that TENS may be used more than 450,000 times in Canadian hospitals each year, with much more use in private clinics. This carries substantial cost implications. The characteristics of TENS use are summarised in *Table 20*, which shows that predominant use is in physiotherapy.

TABLE 20 Use of TENS in Canada

	Number of hospitals	%
Users		
Physiotherapists	43	98
Nurses	7	16
Physicians	2	5
Others	5	11
Departments		
Physiotherapy	39	89
Rehabilitation medicine	6	14
Pain clinics	2	5
Labour and delivery	3	7
Application		
Chronic pain	42	96
Acute pain	41	93
Labour and delivery	19	43
All three	16	36

There is conflicting professional opinion about the use of TENS for acute postoperative pain – although it is used for acute pain in 93% of Canadian hospitals. The recommendations of the Agency for Health Care Policy and Research¹¹⁶ for acute pain management state that TENS is “effective in reducing pain and improving physical function”, while the earlier report of the UK College of Anaesthetists’ working party on pain after surgery¹¹⁷ says that “TENS is not effective as

the sole treatment of moderate or severe pain after surgery”. For postoperative pain some textbooks recommend or strongly recommend TENS,^{118–122} although at least one is uncertain.¹²³

TENS is of doubtful benefit in labour pain,¹²⁴ but we could find no systematic review of its use in chronic pain other than that of Reeve and colleagues.¹¹⁵ This review examined the use of TENS in acute, labour and chronic pain, and concluded that TENS has not undergone sufficiently strict and rigorous clinical evaluation – especially what the authors call the ‘purifying heat’ of RCTs.

The quality of methods used in clinical trials has been shown to be a key determinant of the eventual results. Schulz and colleagues³⁰ have demonstrated that trials that are not randomised or 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%. TENS can almost never be properly blinded,¹²⁵ and bias is likely from this source alone.

For an analgesic drug, evidence of its effectiveness is sought first in standardised acute pain settings and, once established there, its use moves from acute to chronic conditions. In this chapter, although this may not be totally appropriate for TENS, the analgesic efficacy of TENS in acute postoperative pain and labour pain is examined first, and then its effectiveness in chronic pain conditions is discussed.

TENS in acute pain

A full report of the use of TENS in acute pain has already been published.⁷³ 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).³ The terms ‘TENS’ and ‘transcutaneous electrical nerve stimulation’ were used in searching, both alone and in combination. Additional reports were identified from the reference lists of retrieved reports, review articles and textbooks.

Inclusion criteria were full journal publication, TENS, and 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 where the number of patients per treatment group was less than ten were excluded.

Two types of control group 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 that could possibly meet the inclusion criteria was read by each reviewer independently and scored for inclusion and quality using a three-item scale.⁷ Reports included were awarded one point for randomisation, a further point if this had been performed correctly, and a third if the number of, and reasons for, withdrawals were given. Reviewers met to consider whether studies were randomised, or if the description of the method of randomisation was adequate.⁷

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 our judgement of overall effectiveness. Reports which had pain measures but were not randomised or inadequately randomised were examined for positive or negative analgesic effectiveness of TENS, using the judgement of their authors.

Acute pain – results

A total of 19 reports were either not RCTs or the method of randomisation was inappropriate. Of

the 19 reports with pain measures 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.

Randomised studies

The randomised studies included information from 786 patients. 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; individual titration of settings took place in six reports. In 14 reports, TENS was compared with sham TENS without batteries, with batteries reversed or with sub-threshold stimulation; the other three compared TENS plus intramuscular opiate with intramuscular opiate alone. Quality scores were generally 1 or 2, out of the 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.

TENS versus sham TENS

Of the 17 RCTs included, 14 compared TENS with sham TENS. Not one found any difference. One of the 14 reported no significant difference between TENS and sham TENS for analgesic consumption,¹²⁶ but did report a statistically significant difference for pain intensity in favour of the active TENS. However, the published results used a one-tailed statistical test which was judged inappropriate.

TENS versus opiate control

Seven of the 17 RCTs included compared opiate plus TENS with opiate alone, four of which also included sham TENS. Of these 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.^{127,128}

Pike¹²⁷ studied 40 patients after total hip replacement. The study had as its main outcome measure the number of pethidine (meperidine) injections given in the first 2 postoperative days and a retrospective global rating. Patients with active TENS had significantly fewer injections on the first postoperative day as well as higher scores on

global rating of treatment. VanderArk and McGrath¹²⁸ recruited 100 patients having abdominal and thoracic surgery over 2 months, and although there was more success with active TENS used for 20 minutes three times per 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 studies in which no benefit was demonstrated.

Adverse events

No report described systematic recording of adverse events, nor were any reported.

TENS in labour pain

A number of different search strategies were used to identify eligible reports in both Medline (1966–95) and the Oxford Pain Relief database (1950–94). The terms ‘TENS’ and ‘transcutaneous electrical nerve stimulation’, ‘labour’ and ‘childbirth’ were used in searching, both alone and in combination, and without language restriction. Additional reports were identified from the reference lists of retrieved reports, review articles and textbooks. 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 where the numbers of patients per treatment group were fewer than ten were excluded.

Each report which could possibly meet the inclusion criteria was read by each reviewer independently and scored for inclusion and quality using a three-item scale which examined randomisation, blinding and withdrawal and drop-outs. An included study could receive a maximum score of 5 and a minimum of 1. When the method of treatment allocation was unconcealed (alternate allocation, for instance), the report was excluded.

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 placement of electrodes was also extracted. Control group design and the use of TENS in these controls was similarly noted, including methods of disabling TENS devices (sham TENS). 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 the report was positive or negative for the analgesic effectiveness of TENS. *Post-hoc* sub-group analysis in the original reports was not considered in our judgement of overall effectiveness.

Labour pain – results

Eight reports involving 712 women were included; 352 women received active TENS and 360 acted as controls. One study used cranial TENS; others used TENS with dorsal or suprapubic stimulation. Seven different TENS devices were used in the eight studies, predominantly with individual titration. Three studies used conventional (no TENS) analgesic administration as the control group. Five studies used disabled TENS instruments (sham TENS) as a control group. Only one¹²⁹ had made sufficiently determined attempts at blinding to be awarded any points for blinding. This study had a score of 4; six studies scored 2, and one scored 1.

A variety of different pain measures was used, but there was no consistent method of measuring pain intensity or relief. Some studies measured suprapubic and back pain separately, and some measured pain at different stages of labour, or different degrees of cervical dilatation. No study recorded any difference in pain scores during labour between TENS and control. Significant analgesic benefit was recorded in only one study,¹²⁹ where postpartum retrospective questions about overall pain relief elicited that 64 of 132 women who had received TENS had moderate relief or better, compared with 49 of 148 women having sham TENS (odds ratio: 1.89, 95% CI: 1.17–3.05; NNT: 6.5, 95% CI: 3.7–25.2).

Additional analgesic interventions were recorded in six reports. In two, the total number of interventions was noted. Bundsen and colleagues¹³⁰ recorded 17 additional analgesic interventions in 11 women receiving normal obstetric analgesic care, compared with 21 additional interventions in 16 women receiving TENS. Nesheim¹³¹ recorded that 35 women with TENS needed 49 analgesic interventions, compared with 63 interventions in 35 women with sham TENS. Four studies gave figures for the number of women who received any other analgesic intervention (*Figure 20*); 81% of 230 women with TENS had an analgesic intervention, compared with 89% of 245 women having sham

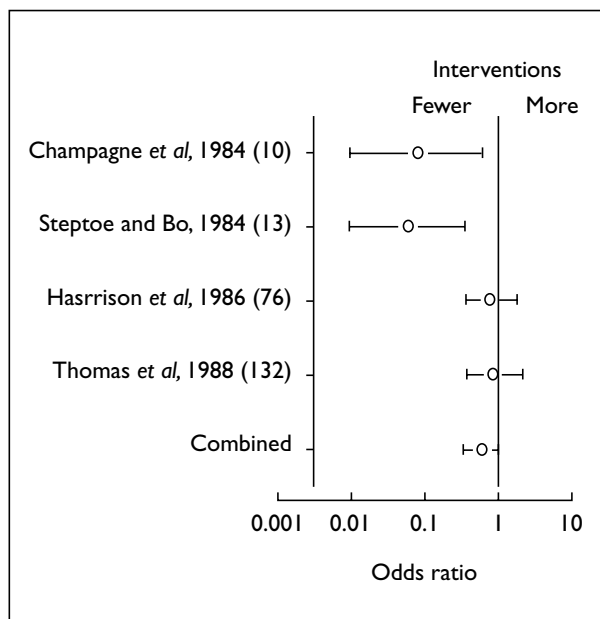


FIGURE 20 Odds ratios for effect of TENS in labour on use of other analgesic interventions

TENS. While the two larger studies^{129,132} were not statistically significant in themselves, the combined results of all four studies had an odds ratio of 0.57 (95% CI, 0.34–0.96). With an NNT of 14 (95% CI, 7.3–119), this meant that 14 women would need to receive TENS for one to benefit by not needing further analgesic intervention.

Of the eight studies, three were judged by us to have a positive result, and five a negative result. The three positive studies included: the study of cranial TENS in 20 women,¹³³ which had additional pain relieving measures as its only outcome; another small study with 25 women,¹³⁴ and a larger study whose positive outcome was increased time both to and between epidural local anaesthetic injections.¹³⁵

There were no reports of adverse events.

TENS in chronic pain

Reports of RCTs for TENS in chronic pain were sought. A number of different search strategies were used to identify eligible published reports in Medline (1966–96) using Knowledge Finder v. 3.3, CINAHL (using MacSPIRS v. 2.3), the Oxford Pain Relief database (1950–94),⁶ and Embase (1980–96). The latest indexing date used was May 1996. The terms ‘TENS’, ‘TNS’ and ‘transcutaneous electrical nerve stimulation’ were used in searching, including both single

words and combinations. Searching was performed using both MeSH terms and free text words. Additional reports were identified from the reference lists of retrieved reports, review articles and textbooks.

Included were full journal publications of RCTs in which the analgesic effects of TENS was studied in patients with chronic pain. Reports where TENS was used under experimental pain conditions, and where no clinical pain outcome measurements were used, were excluded.

Each report which could possibly meet the inclusion criteria was read by each reviewer independently and scored for inclusion and quality using the three item scale.⁷

Several types of controls were found, sham TENS (no current, distant control points), active interventions (e.g. oral NSAID), or various combinations of these. In some cases, TENS was compared with TENS (high versus low frequency stimulation). Some studies used blinded observers. While there was no prior hypothesis that TENS cannot adequately be blinded before the original reports were read, it was determined at the consensus meeting that, despite the considerable efforts documented in some reports, adequate blinding was practically impossible. Consequently, no report was given any points for blinding, even if they were described as double-blind or used blinded observers. Thus, an included report could have a maximum score of 3 and a minimum score of 1.

Information about the pain condition, numbers of patients, study design and year of publication was extracted from the reports. Type of TENS equipment, duration of each treatment period, hours of TENS per week, and numbers of TENS sessions were recorded. High frequency TENS (> 10 Hz), low frequency TENS (< 10 Hz), continuous or pulse stimulation, intensity of stimulation above or below the sensory threshold, and sites of electrode were identified. 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 or drop-out information.

A judgement was then made as to whether the overall conclusion of the report was positive or negative for the analgesic effectiveness of TENS. *Post-hoc* sub-group analysis in the original reports was not considered in our judgement of overall effectiveness.

Chronic pain – results

In all, 38 RCTs on chronic pain were included. TENS exposure in these trials was low: duration of treatment was less than 4 weeks in 83% of the trials. In 85% of trials, stimulation was for less than 10 hours per week and 67% of the patients had less than ten TENS sessions in their trial.

Ten out of 24 trials comparing TENS with a control (no current, placebo tablets or control points) were regarded having a positive outcome.

In three out of 15 trials comparing TENS with active treatment, a positive outcome was noted.

In ten reports, a comparison was made between frequency and mode. In four out of five trials, conventional high frequency TENS equalled low frequency pulsed TENS.

Comment

Methods used in trials of TENS

The 'gold standard' in clinical trials is adequate randomisation.³⁰ Non-randomised studies have, for nearly 20 years, been shown to yield larger estimates of treatment effects than studies using random allocation.¹³⁶ The degree of the exaggeration of treatment effect when randomisation is inappropriate can be as much as 40%.³⁰ These findings underpin the inclusion criteria chosen in systematic reviews.

For TENS in acute postoperative pain, for example, 17/19 reports with pain outcomes, which were either not randomised or inappropriately randomised, claimed TENS to be effective, compared with 2/17 RCTs.

The possibility of bias exists in those trials described as randomised. The method of randomisation was described in only two reports.^{137,138} The method described was inadequate in both, one using a nurse to randomise patients¹³⁷ and the other using alternate allocation.¹³⁸ Reports which said only that they were randomised may have used an inadequate method.

While randomisation should exclude selection bias in trials, blinding should exclude observer bias. Inadequacy of blinding in clinical trials of analgesic interventions continues to be of concern,⁴¹ although this may be less of an issue with pharmacological interventions.⁶⁶ 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 the allocation to conduct the patient assessments. Lack of blinding has been estimated to exaggerate the estimate of treatment effect of trials by some 17%.³⁰ Adequate blinding of TENS for both carers and patients is particularly difficult.¹²⁵

None of the reports of TENS in acute pain 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. For TENS in labour, three studies made no attempt at blinding and, of the five that used sham TENS, only one¹²⁹ described the method of blinding in sufficient detail to indicate that blinding may have been adequate. It is extremely difficult to exclude observer bias in trials of TENS,¹²⁵ and results must be so judged.

Classic analgesic trial methods study patients given placebo and those given a standard analgesic to establish that the study has sensitivity – that it can actually measure an analgesic response. An analgesic response with a new treatment measured in the same study then has validity. This approach has been found important because group sizes in analgesic trials are usually of less than 50 patients, so that small differences in the number of responders in placebo or analgesic groups can affect results profoundly. Studies with an A versus B design, such as all those identified with TENS in acute pain and labour, do not have measures of internal sensitivity and this makes the results more difficult to interpret, especially where comparators are not known in themselves to be effective (as with comparisons of TENS and sham TENS).

The choice of outcome measure is also an important determinant of how studies are to be judged. TENS is considered to produce pain relief. The most important outcome should therefore be lower pain intensity or greater pain relief. Use of other analgesic interventions is a secondary outcome measure but one commonly used in these studies. Choosing TENS for any future labour, for example, is an even less direct measure of analgesic effectiveness.

What we are left with, therefore, is the lowest common denominator of information, essentially vote counting rather than a more sophisticated

analysis, reflecting 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 statistic⁶⁶), or mean analgesic consumption, or time to first analgesic was used. None of these allowed data extraction for further statistical analysis or 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.

TENS in acute pain

For TENS in acute pain, there is abundant evidence of lack of analgesic effect. Of 17 trials, 15 could not show a clear analgesic effect – and those trials were themselves subject to bias from lack of adequate blinding and possibly from lack of adequate concealment of randomisation.

TENS in labour pain

The overwhelming conclusion for the use of TENS in labour is also that there is evidence of lack of effect. The only report which produced any difference between TENS and control for any pain score was that of Thomas and colleagues.¹²⁹ When labour was over, the patients were asked how much pain relief they felt they had received from TENS; 48% of women with TENS felt they had moderate, good or excellent relief compared with 32% of those having sham TENS, an odds ratio of 1.89 (95% CI, 1.17–3.05) and an NNT of 6.5 (95% CI, 3.7–25.2). This is the best evidence for an analgesic effect of TENS in a study with large numbers of women and with high methodological quality, although the study found no differences in pain scores collected during labour between active and sham TENS.

The secondary measure of additional analgesic interventions also showed possible benefit from TENS. The magnitude of the effect was not great – about 81% of women having TENS had an additional analgesic intervention compared with 89% with sham TENS controls. While this reduction in analgesic co-interventions was statistically significant for four studies combined, the two larger studies showed no significant difference. Of 14 women using TENS during labour, one woman will not have an analgesic intervention who would have done had TENS not been used (NNT, approximately 14). Confidence in that estimate is not great, with a lower 95% CI of 7, and a higher CI of 119 women needing to use TENS for one to be spared an analgesic intervention. Where epidural local anaesthetics are being given, TENS may delay injections by perhaps 10 or 20 minutes.

Where blinding is equivocal, these minor benefits must be questioned. Since lack of blinding or inadequate blinding may give rise to an overestimation of treatment effect by almost 20%, the small analgesic-sparing effect of TENS in labour may be due to observer bias.

Thomas and colleagues¹²⁹ asked women about the use of TENS in future labours. In the group that had received TENS, 54% claimed that they would choose TENS again, compared with 32% in the sham TENS control group, a statistically significant difference.

The evidence from RCTs for analgesic benefits from TENS during labour is not compelling. Though there are tantalising hints of analgesic-sparing effects, and while women seem to appreciate some features of the use of TENS during labour by tending to favour it for future births, it is difficult to ignore the possibility that all of these may be the result of some observer bias, or observer–patient interaction. The argument may be that something that appears to do no harm, and which may do some good, is worth using. There is a cost to that, both in use of resources for little hard benefit, and in diverting attention away from a research agenda which should perhaps be more sharply focused on the means of providing safe and effective analgesia for women in labour.

If there is any research agenda for the use of TENS during labour, then what is needed is a very large randomised, open study which addresses effects on clinically relevant outcomes, especially any sparing effect of TENS on more invasive, and potentially more harmful interventions, such as the use of epidural local anaesthetics.¹³⁹

TENS in chronic pain

For TENS in chronic pain, the situation is different. Here we have lack of evidence of effect rather than evidence of lack of effect. The issue is one of dose. Many, perhaps most, chronic pain physicians who use TENS prescribe at least 30 minutes use twice a day for at least a month before any effect may be felt.¹⁴⁰ This pragmatism is supported by the important study of Nash and colleagues,³³⁶ who demonstrated a clear improvement in analgesic effects of TENS in a large number of chronic pain patients over a long period.

None of the RCTs used doses of TENS which approached this. Duration of treatment was less than 4 weeks in 83% of the trials, and in 85% of the trials stimulation occurred less than 10 hours per week, with 67% of the patients having less than ten

sessions of TENS. Even anecdotal reports of the success of TENS suggest that dose is important; a report in *The Times* by Thomas Stuttaford quoted Grahame Le Saux (a Blackburn footballer) as using TENS for at least 30 minutes before and after every training session and match to achieve benefit, as well as at other times.

The use of TENS in chronic pain may well be justified but it has not been seen. There is a requirement for a randomised trial to address the issue. It will be difficult to design and organise, it will need to be multicentred in the UK and other European countries may need to be included, it will require large numbers of patients and simple outcome measures. Without it, a potentially valuable intervention may be

underused, or a useless intervention may continue in use.

Conclusion

TENS is of no value in acute pain.

TENS is of possible value in labour pain because it may 'spare' other analgesic interventions which carry increased morbidity for mother and baby. A large randomised trial is needed to prove this.

TENS may be useful in chronic pain but there is no useful evidence. A large, multicentre, randomised trial of TENS with sufficient dose and duration is needed.

Chapter 9

Relaxation for relief of chronic pain

A systematic review of RCTs was undertaken to examine the effectiveness of relaxation therapy in chronic pain (Carroll & Seers, in press).

A number of different methods were used to identify eligible reports. These included searching of the following electronic databases using both Knowledge Server v. 3.25 and WINSPIRS v. 2.32 as the search platforms: Medline (1966–6/1996), Psychlit (1974–6/1996), CINAHL (1982–6/1996), Embase (1980–6/96) and the Oxford Pain Relief database (1950–6/1994).⁶ Studies were included if they were RCTs of relaxation techniques in chronic pain. Studies which investigated the effects of relaxation in combination with other interventions were not considered.

Nine studies of 414 patients met the pre-defined inclusion criteria and are critically appraised in this qualitative review; 196 patients received relaxation. Sample sizes for the relaxation-only groups ranged from 12 to 30. Meta-analysis was not possible because of lack of quantitative data in the primary studies. Studies were in a range of chronic pain conditions including low back pain (2), rheumatoid arthritis (1), temporo-mandibular joint dysfunction (1), fibromyalgia (1), ulcerative colitis (1), and cancer (2). The most common form of relaxation was progressive muscle relaxation.⁶ Two studies used tapes. The McGill Pain Questionnaire was the most common pain outcome used.

Whilst many studies were able to show a significant difference between the pre- and post-test assessments for the pain outcomes used, few statistically significant differences were reported in favour of relaxation on between treatment comparisons. Only three studies reported statistically significant differences in favour of relaxation (judged as a significant difference for at least one of the pain outcomes used) compared to the other treatments groups. In rheumatoid arthritis, the McGill Pain Questionnaire scores were significantly lower for patients receiving relaxation compared with those who were in the routine treatment control group.

In ulcerative colitis, significant differences were reported for six of seven different pain outcome measures in favour of progressive muscle relaxation compared with patients on the waiting list control group.

In one of the two cancer pain studies, relaxation taught by nurses produced significantly lower pain sensation scores compared with the control group.

Two studies reported significant differences in favour of the experimental control groups rather than for relaxation.

There is insufficient evidence to confirm the analgesic effectiveness of relaxation in chronic pain. Many of the positive and negative studies suffer methodological inadequacies.

Chapter 10

Intravenous regional sympathetic blockade for reflex sympathetic dystrophy

Summary

This systematic review of intravenous regional sympathetic blocks (IRSB) in patients with reflex sympathetic dystrophy (RSD) included eight RCTs. Six used guanethidine; none showed significant analgesic effect in IRSBs to relieve pain due to RSD. Two reports, one using ketanserin and one bretylium with 17 patients in total, showed some advantage of IRSBs over controls. RCT results were not combined because of the variety of different drugs and outcome measures and because of methodological deficiencies in most of the reports.

Our own trial was stopped prematurely because of the severity of the adverse effects. No significant difference was found between guanethidine sulphate and placebo on any of the outcome measures. Patients in all groups reported less than 30% of the maximum possible relief during the first week after the injections and on only two occasions (one with saline and one with low dose guanethidine) was relief reported for longer than a week. There was no evidence of a dose response for guanethidine.

The use of guanethidine in IRSBs for patients with RSD was not supported by the systematic review. The complete version of this study is published elsewhere.¹⁴¹

Introduction

Over 120 years ago, Weir Mitchell described a syndrome of persistent burning pain and trophic changes in the limbs of soldiers after gunshot wounds; he called it *causalgia*. Since then several similar syndromes have been described, using different names. The term 'reflex sympathetic dystrophy' was first used over 40 years ago, and today it is used to describe the wide range of chronic pain conditions associated with hyperactivity of the sympathetic nervous system.^{142,143} There is a difficulty with this definition – the implication that it is a disordered sympathetic nervous system which

causes the pain; it may be that the pain causes a disorder of the sympathetic nervous system. The clinical reality is a funny pain in a funny-looking limb. The pain is usually constant, severe and unresponsive to conventional analgesics. It is usually confined to a limb, and the limb may show hyperaesthesia, swelling, changes in skin colour and temperature, changes in sweating and even bone demineralisation.¹⁴⁴

The lack of response of these pains to conventional analgesics led to the development of techniques designed specifically to block the sympathetic nervous system. One of these techniques, described more than 15 years ago,¹⁴⁵ is IRSB, which involves giving a drug known to block the sympathetic nervous system (guanethidine in the original report) in high local concentration in a limb isolated with a tourniquet. Since it was first described, this technique has gained considerable popularity, mainly because of its simplicity and relatively low cost. However, even though IRSB is recommended as the simplest, most effective and safest way to relieve pain associated with sympathetic hyperactivity,^{146,147} very few RCTs have assessed its effectiveness.

We undertook a systematic review of the literature for IRSBs in patients with RSD and, because of the paucity of controlled data supporting the use of guanethidine in IRSBs, performed a double-blind crossover randomised study. This was designed to assess the effectiveness of IRSBs with guanethidine, in patients with RSD who had claimed relief after open blocks, and to determine whether the analgesic effects (if any) were dose-dependent.

Systematic review

Methods

A Medline (Silver Platter Medline, v. 3.0 and 3.1) search was undertaken from 1966 to May 1993. The strategy was designed to identify the maximum number of randomised and/or double-blind studies by using a combination of text words, 'wild cards' and MeSH terms as described previously.³

Medical journals were hand-searched. They were selected from a list of the 50 journals with the highest number of reports in Medline using the optimised strategy, and nine specialist journals which were not included in that list or were not indexed in Medline. The search covered volumes published between 1950 and 1992, and the order in which journals were searched was determined mainly by local availability.

The eligibility of a study was determined by looking only at the methods section. The following criteria had to be met:

- (i) inclusion of patients with chronic pain associated with RSD
- (ii) random and/or double-blind allocation of treatments
- (iii) administration of an IRSB to at least one of the treatment groups.

A letter was sent to the first author of each eligible study requesting individual data in relation to analgesic measures and adverse effects for all the patients included in the study as well as drop-outs. The authors were asked to describe the randomisation method used in the study and to send details of any other trials fulfilling our inclusion criteria performed by them or by other research groups.

Results

Eight controlled trials were found that assessed the analgesic effects of IRSBs^{141,148-154} in patients with RSD (*Table 21*). Sample sizes ranged from six to 21 patients.

In most reports guanethidine was the active substance, but reserpine, bretylium tosylate, droperidol and ketanserine were used as alternatives.

TABLE 21 RCTs of IRSB in RSD

Trial	No of patients	Design [†]	Experimental group(s)	Control group	Treatments per group	Outcomes	Results [‡]
Bonelli <i>et al</i> , 1983 ¹⁴⁸	19	P, O	Guanethidine, 20 mg, + heparin, 500 U	Stellate ganglion block with bupivacaine, 75 mg	4 experimental 8 control	Mean daily pain intensity	NS Hyperpathia: 2/7 controls & 4/7 IRSB improved after treatment (no follow-up).
Kettler <i>et al</i> , 1988 ¹⁴⁹	6	C, DB	Droperidol, 2.5 mg, + heparin, 500–1000 U	Heparin, 500–1000 U, in normal saline	1	Description of current and daily VAS pain intensity	NA On three occasions relief with control, on three relief with droperidol. Only sustained (2 weeks) relief after placebo.
Rocco <i>et al</i> , 1989 ¹⁵¹	10	C, DB	Guanethidine, 20 mg, or reserpine, 1.25 mg, + lignocaine, 250 mg	Lignocaine, 250 mg	Variable	Mean pain intensity	NS Pain relief reported for 2 months or more by 2/12 on reserpine, 1/12 on guanethidine.
Hanna <i>et al</i> , 1989 ¹⁵⁰	9	C, DB	Ketanserine, 10–20 mg	Normal saline	2	Mean weekly pain intensity	S No individual data.
Blanchard <i>et al</i> , 1990 ¹⁵²	21	C, DB	Guanethidine, 20–30 mg or reserpine, 0.5–1.0 mg	Normal saline	1	Reduction in VAS pain intensity > 50%	NS After 4 weeks 0/21 had relief with saline, 1/21 with reserpine, 3/21 with guanethidine.
[†] P, parallel; C, crossover; DB, double-blind; O, open. [‡] S, significant; NS, not significant; NA, not applicable (statistical analysis was not possible due to the number of withdrawals).							
							<i>continued</i>

TABLE 21 contd RCTs of IRSB in RSD

Trial	No of patients	Design†	Experimental group(s)	Control group	Treatments per group	Outcomes	Results‡
Hord <i>et al</i> , 1991 ¹⁵³	8	C, DB	Bretylum, 1.5 mg/kg, + lignocaine, 200–300 mg	Lignocaine, 200–300 mg	2	Duration of VAS relief > 30 mm	S No individual data.
Dhar <i>et al</i> , 1992 ¹⁵⁴	15	C, DB	Guanethidine, 15 mg, + lignocaine 1%, 10 ml	Lignocaine, 1% 10 ml	1	Somato-sensory tests + VAS pain intensity (provoked and on-going)	NS No individual data.
Jadad <i>et al</i> , 1995 ¹⁴¹	8	C, DB	Guanethidine, 10 and 30 mg in saline 25 ml (arm) OR 20 and 30 mg (leg) in saline 50 ml	Saline 25 ml (arm) or 50 ml (leg)	1	VAS relief, pain diaries	NS Study stopped because of hypotension in two patients.

† P, parallel; C, crossover; DB, double-blind; O, open.
‡ S, significant; NS, not significant; NA, not applicable (statistical analysis was not possible due to the number of withdrawals).

Of the eight RCTs, two (17 patients in total) showed some advantage of IRSBs over control treatments.^{150,153} Neither study used guanethidine. The remaining trials failed to show a significant difference in analgesic response between the IRSBs and the control treatments. Only one¹⁴¹ of the five guanethidine trials was designed to determine if there was a dose-response for the effects of guanethidine.

The analgesic outcome was measured in many different ways, so that it was not possible to combine the results mathematically. Most studies had serious methodological deficiencies. The major methodological problems of the studies were poorly defined diagnostic criteria,^{150,152,153} inadequate washout periods,¹⁵⁰ incomplete cross-over,^{151,153} open administration of treatments,¹⁴⁸ no description of the technique,¹⁵⁴ and a high proportion of withdrawals with no 'intention-to-treat' analysis.^{149,153}

There was no RCT evidence to support guanethidine use in IRSBs to treat RSD, but some evidence for ketanserin and bretylum. All the studies were small, so that there is substantial risk of false-negatives and even false-positives.

Comment

Persistent pain due to RSD is usually difficult to control and can be very frustrating for both patients and doctors. Although the most widely used management is IRSB, the results of the few

RCTs available are conflicting. Two small studies, neither using guanethidine, did show significant relief.^{150,153} There is thus some limited support for the technique but none for guanethidine as the active drug in the injection to relieve the pain associated with RSD. Only one RCT has ever shown analgesic effects of blocks with guanethidine.¹⁵⁵ It was not included in the systematic review because the patients in the study had rheumatoid arthritis, not RSD.

In our own study, two patients suffered hypotension so the study was halted prematurely.¹⁴¹ The technique is not innocuous. Prolonged arterial hypotension has been described after multiple IRSBs¹⁵⁶ but, in our experience, it can occur after a single injection. Another important lesson is that some patients with RSD do not respond to IRSB. There are no known predictors of poor response. The argument that patients who do not respond to a sympathetic block do not have RSD is a rather circular one, particularly since the effectiveness of IRSBs for pain relief in RSD has not been proven. It remains unclear whether the sympathetic hyperactivity is the cause or a result of the pain, or a confounding factor. The poor analgesic response reported by patients in these trials makes it difficult to justify the use of this technique using guanethidine for the treatment of RSD. The defence, that satisfactory results may follow multiple injections,^{145,147} is not supported by double-blind RCTs.

Other drugs given as IRSBs may produce pain relief for RSD. Two RCTs (17 patients in total) of IRSBs

using ketanserin and bretylium showed significant relief in patients compared with controls. The technique itself may also produce pain relief, although our results provided little evidence of this. One hypothesis is that pain relief after IRSB could be due to ischaemia induced by the tourniquet

rather than the injection.^{152,157} If this is so, it would be much simpler and safer to use the tourniquet as the treatment; this could be done by the patient at home. This would fail if the pain relief is due to the placebo effect of the procedure (including intravenous injections) performed in the hospital.

Chapter 11

Epidural corticosteroids for sciatica

Two systematic reviews have addressed the effectiveness of epidural steroid injections for sciatica and back pain. Both have examined the RCTs published up to the end of 1994.

The analysis by Koes and colleagues from Amsterdam¹⁵⁸ goes into great depth examining methodological quality of the trials and how this has been scored by the reviewers. This type of review is, frankly, disappointing; it does not make any meta-analytical judgements on which to work and produces few enlightening ideas about a future research agenda, apart from some anodyne comments about possible trial design.

The best that this review can do is to state that, of the four studies with the highest methodological quality assessed by their particular scoring system, two had positive outcomes for epidural steroids (judged by their authors) and two had negative outcomes.

In contrast, a meta-analysis of the same trials by Watts and Silagy¹⁵⁹ is an important step forward in showing that epidural corticosteroids do have an analgesic effect on sciatica compared with controls. Their analysis, using odds ratios, answered the question, 'Do epidural steroids work?'

We wished to address the question, 'How well do they work?' to try to assess the extent of the benefit given by the steroids. In order to do this, we re-analysed their data making NNT the measure of clinical benefit and using the outcome of at least 75% pain relief for short-term outcomes (1–60 days) and at least 50% pain relief for long-term outcomes (12 weeks–1 year). In circumstances where the numbers of patients entered into individual studies is small, this approach is most likely to produce a reliable indication of the clinically-relevant outcomes needed by patients, providers and purchasers in making decisions about the use of a potentially harmful technique. Evidence from other sources places the risk of neurological sequelae after an epidural as 1 in 5000.¹⁶⁰

The NNT for short-term (1–60 days), greater than 75% pain relief from the ten trials with short-term outcomes combined, was just

under 6 (95% CI, 4–12). This means that of six patients treated with epidural steroid, one will obtain more than 75% pain relief in the short term who would not have done had the control treatment (placebo or local anaesthetic) been prescribed (*Table 22* and *Figure 21*).

The NNT for long-term (12 weeks–1 year) improvement from the five trials combined, was about 11 for 50% pain relief (95% CI, 6–90). This means that of 11 patients treated with epidural steroid, one will obtain more pain relief over this longer period who would not have done had the control treatment (placebo or local anaesthetic) been prescribed (see *Table 22* and *Figure 22*).

These NNT values appear, at first sight, to be disappointing. Here is an intervention which shows statistically significant improvement compared with controls, and yet the clinical benefit, the NNT for one patient to reach the chosen end-point, is 6 for short-term benefit and 11 for long-term. The short-term end-point chosen, however, is quite a high hurdle. Using the easier hurdle of 50% pain relief rather than 75%, the 'best' NNT achieved by drug treatment of neuropathic pain was just under 3. Patients may choose an epidural if it means they do not have to take medication and, particularly, if it gives a higher level of relief, even though there is only a 1 in 6 chance of this level of response.

The long-term NNT of 11 is, perhaps, not surprising. Occasional patients in most clinics report a 'cure' as a result of a steroid epidural but the majority of epidural steroid 'successes' return for repeat epidurals. The fact that one patient experiences relief lasting between 12 weeks and 1 year for 12 patients treated with epidural steroid fits with our experience.

The message is that we will have inevitably to expose our practice to the searching type of analysis which Watts and Silagy¹⁵⁹ have used for epidural steroid. This intervention has shown a statistically significant benefit over control. Other interventions will not show this benefit and will be discarded. For those interventions which do show statistically significant benefit over controls there is then a further stage, which is to define the clinical benefit of the intervention. The NNTs for

TABLE 22 Epidural corticosteroids for sciatica

Trial [†]	Number of patients		Relative risk (95% CI)	NNT (95% CI)
	Improved on epidural steroid	Improved on control		
Bush, 1991	10/12	7/11	1.3 (0.8–2.2)	5.0 (1.8–∞)
Cuckler, 1985	11/46	4/27	1.6 (0.6–4.6)	10.99 (3.66–∞)
Dilke, 1973	16/43	8/38	1.8 (0.9–3.7)	6.3 (2.8–∞)
Mathews, 1987	9/23	14/34	0.95 (0.5–1.8)	–
Swerdlow, 1970	76/117	98/208	1.4 (1.1–1.7)	5.6 (3.5–14)
Combined long-term	122/241	131/318	1.2 (1.03–1.47)	11 (5.6–90)
Beliveau, 1971	18/24	16/24	1.1 (0.8–1.6)	12 (3–∞)
Breivik, 1976	9/16	5/19	2.1 (0.9–5.1)	3.3 (1.6–∞)
Bush, 1991	8/12	2/11	3.7 (0.98–13.7)	2.1 (1.2–7.5)
Cuckler, 1985	12/42	8/31	1.1 (0.5–2.4)	33 (4–∞)
Dilke, 1973	21/35	11/36	2.0 (1.1–3.4)	3.5 (1.9–14)
Klenerman, 1984	15/19	32/44	1.1 (0.8–1.5)	17 (3–∞)
Mathews, 1987	14/21	18/32	1.2 (0.8–1.8)	10 (3–∞)
Popiolek, 1991	18/28	8/30	2.4 (1.3–4.6)	2.6 (1.6–7.2)
Ridley, 1988	17/19	3/16	4.8 (1.7–13)	1.4 (1.1–2.1)
Snoek, 1977	8/27	5/24	1.4 (0.5–3.8)	11 (3–∞)
Combined short-term (> 75%)	140/243	108/267	1.4 (1.2–1.7)	5.9 (3.9–11)

References to the original papers are given in Watts and Silagy.¹⁵⁹ Relative risk and NNT were calculated.

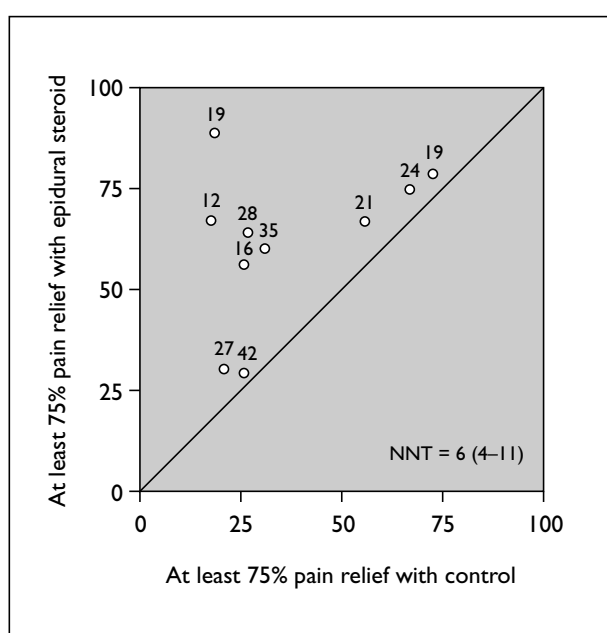


FIGURE 21 Epidural steroids for sciatica: short-term relief (1–60 days)

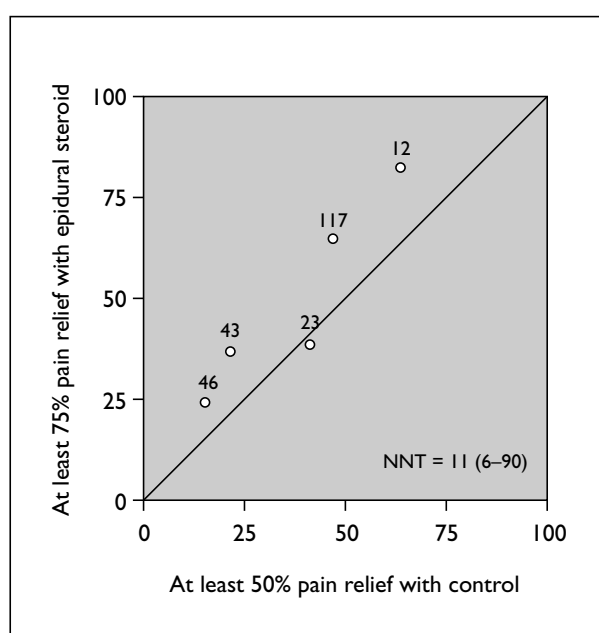


FIGURE 22 Epidural steroids for sciatica: long-term relief (12 weeks – 1 year)

effectiveness are one possible definition, particularly when coupled with NNTs for minor and major harm (for example, the NNT for dural tap of about 40¹⁵⁹).

For patients with chronic disease, and in this case painful chronic disease, interventions may be attractive even if their success rate is far lower than would be acceptable in, say, the management of postoperative pain. This means that the interpretation of measures of clinical benefit, such as

NNTs, has to be context-dependent. For the moment, as Watts and Silagy have demonstrated,¹⁵⁹ we need the best possible analysis of the data available. If the data are poor then the clinical research agenda is established. If the data are reasonable then an attempt to define measures of clinical benefit can be made. The art of clinical practice will then come into play, as patient and doctor juggle the risks and benefits of the alternatives, albeit with better data than presently available.

Chapter 12

Spinal cord stimulation for back pain

Spinal cord stimulation (SCS) for chronic pain, like TENS, is based on the gate control theory of pain. Surgically implanted electrodes (usually but not necessarily over the dorsal columns) are stimulated to activate pain inhibitory mechanisms. The number and type of electrodes implanted, and the type of stimulation received, is very variable. Implantation can be at open laminectomy, or may be percutaneous. The neurophysiology is unclear.

Turner and colleagues¹⁶¹ have conducted a thorough literature search in an attempt to determine the place of SCS in chronic pain treatment. Their literature search was conducted up to June 1994 and the paper was published in December 1995. They included 39 studies but none were randomised – all were case series. A majority (82%) did not have planned study protocols. The review does not give the total number of patients treated or reported. Few studies defined patient inclusion or exclusion criteria, or gave demographic information on patients treated. The source of follow-up data was unclear in most of the studies. Most studies reported on percutaneous electrode implant with single-channel stimulators.

In all, 29 studies had at least 50% pain relief as an outcome. Across these studies the mean (probably unweighted, but not stated) was 59% of patients achieving this outcome at some follow-up point, with a range of 15–100%. In 14 studies, 1-year follow-up was reported, with success defined as at least 50% pain relief with stimulator in use. In these, the mean across-study success rate was 62%, again with a range of 15–100%.

Fewer studies reported success at later follow-up times. At 2 years, five trials reported a mean of 64% success (range, 55–74%). At 5 years, three studies

reported a mean of 53% success (range, 50–55%). At 10 years, one study reported 35% success.

Complications were common. In 13 studies there was at least one complication in 42% of patients (range, 20–75%). These were predominantly stimulator or electrode problems (mean, 30% and 24%, respectively; range, 0–75%). Infection was less common, occurring in 5% of patients in the 20 trials that reported it (range, 0–12%). Most complications appeared to be minor.

Turner and colleagues¹⁶¹ also reported on an on-going randomised trial, but with limited information and none on pain relief.

It is clear that the evidence on SCS is limited. It is likely to be biased because of lack of randomisation and blinding, because it has been conducted by enthusiasts, and because reporting has been limited. For example, more than half of the studies did not report the numbers of patients who had received implants during the study period.

Turner and colleagues call for a randomised trial or trials with common design, including:

- randomisation to SCS or control
- follow-up assessment of all patients at uniform times after implantation or entry – 6 months, and yearly thereafter
- follow-up assessment by independent observer
- description of all relevant clinical and demographic information
- use of common and valid pain and other measures
- assessment of multiple outcomes, including back pain, leg pain, physical functioning, drug use, work status, healthcare use and quality of life.

Chapter 13

Steroid injections for shoulder disorders

A systematic review of the use of steroid injections for shoulder disorders¹⁶² indicates that the evidence for the efficacy of these is sparse, and what there is shows poor effect.

The reviewers claim that 10% of the total population have one or more episodes of shoulder pain and/or stiffness during their life, and that 5% of all primary care consultations are about shoulder problems. They estimate that 23% of new consultations resolve within 1 month, 51% within 6 months and 59% within 1 year.

They suggest that, in The Netherlands, 12% of all patient-physician contacts for shoulder disorders involve local steroid injections and that injection therapy is given in 20% of all episodes of shoulder disorders. They also indicate that 40% of the patients still had their shoulder problem after 1 year, which means that a considerable proportion of these (common) shoulder problems are not self-limited. These figures seem high for UK practice.

The reviewers found 22 studies that met their inclusion criteria. They put these studies through their tough quality scoring system. No study scored more than 60 out of the maximum of 100, with only three scoring more than 50. This is a clear signal that definitive conclusions are unlikely to be possible from the available studies.

A further complication was that the studies looked at many different treatments, not just at injections.

Only three studies compared steroid injection with saline injection, and five compared steroid injection with injection of local anaesthetic. Yet another difficulty was that the studies used different outcome measures.

Taking the crude criterion of success at 4 weeks or later after injection, the NNT for such success with steroid injection compared with saline injection (three studies) was 17, with a confidence limit that includes no benefit to any patient. For steroid plus local anaesthetic injection versus local anaesthetic alone (five studies), the NNT for success at 4 weeks or later was 33, again with a confidence limit that includes no benefit to any patient. On this basis, one patient in 17 would achieve 'success' with a steroid injection compared with a saline injection, and one patient in 33 would achieve 'success' with a steroid plus local anaesthetic injection compared with an injection of local anaesthetic alone.

Comment

The evidence that steroid injections for shoulder problems are worthwhile is less than compelling. The onus is on those who wish to continue to offer steroid injections to the shoulder to produce convincing evidence. A starting point would be to use a study design that scored closer to the maximum on the Dutch scale.

Chapter 14

Anticonvulsant drugs

Summary

This systematic review of the effectiveness and adverse effects of anticonvulsant drugs used in chronic pain included RCTs identified by Medline, by hand-searching, by searching reference lists and by contacting investigators. NNTs were calculated from dichotomous data for effectiveness, adverse effects and drug-related study withdrawal, for individual studies and for pooled data.

In all, twenty RCTs of four anticonvulsants were considered eligible. Three placebo-controlled studies of carbamazepine in trigeminal neuralgia had a combined NNT for effectiveness of 2.6, for adverse effects 3.4, and for severe effects (withdrawal from study) 24. Three placebo-controlled studies of diabetic neuropathy had a combined NNT for effectiveness of 3, for adverse effects 2.5, and for severe effects 20. Three placebo-controlled studies of migraine prophylaxis had a combined NNT for effectiveness of 2.4, for adverse effects 2.4, and for severe effects 39.

Phenytoin had no effect in irritable bowel syndrome, and carbamazepine little effect in post-stroke pain. Clonazepam was effective in one study of temporomandibular joint dysfunction.

Although anticonvulsants are widely used in chronic pain relief, surprisingly few RCTs show analgesic effectiveness. No RCT compared different anticonvulsants.

Introduction

Anticonvulsant drugs have been used in pain management since the 1960s, very soon after they were first used in medicine and revolutionised the medical management of epilepsy. The clinical impression is that they are useful for neuropathic pain, especially when the pain is lancinating or burning.¹⁶³ Although these disorders are not common (the incidence of trigeminal neuralgia is 4/100,000 per year¹⁶⁴), they can be very disabling. Carbamazepine is one of few effective interventions for trigeminal neuralgia, for which it is usually the drug of choice.¹⁶⁵ In the UK, carbamazepine is licensed for paroxysmal pain of trigeminal

neuralgia (up to 1600 mg daily); phenytoin is also licensed as second-line therapy to carbamazepine in trigeminal neuralgia, either if carbamazepine is ineffective or in patients who cannot tolerate effective doses. When anticonvulsants are used as adjuvant drugs in other pain syndromes, sodium valproate is often preferred to carbamazepine because it may be better tolerated.¹⁶⁶ Anticonvulsants are also prescribed in combination with antidepressants, for example, in the treatment of postherpetic neuralgia.¹⁶⁷ In the UK, no anticonvulsant is licensed for any pain indication other than trigeminal neuralgia.

The precise mechanisms of action of anticonvulsant drugs remain uncertain. There are two standard hypotheses: enhanced inhibition of gamma-aminobutyric acid (valproate, clonazepam), or a stabilising effect on neuronal cell membranes. A third hypothesis is action via N-methyl-D-aspartate receptor sites.

Anticonvulsant drug use is not without risk: serious effects have been reported, including deaths from haematological reactions.¹⁶⁸ The commonest adverse effects are impaired mental and motor function, which may limit clinical use, particularly in the elderly.¹⁶⁸⁻¹⁷⁰

The purpose of this review is to evaluate the analgesic effectiveness of anticonvulsant drugs in order to provide evidence-based recommendations for clinical practice.

Methods

Reports included in this review were RCTs which investigated the analgesic effects of anticonvulsant drugs in patients, with pain assessment as either the primary or a secondary outcome. Excluded were studies that were non-randomised, studies of experimental pain, case reports, clinical observations and studies of anticonvulsants used to treat pain produced by other drugs.

Reports were identified by several methods. A Medline search (Silver Platter v. 3.0, 3.1 and 3.11) from 1966 to February 1994 was undertaken using a search strategy designed to identify the maximum

number of randomised and/or double-blind reports using a combination of text words, 'wild cards' and MeSH terms.³ This search strategy was narrowed to include specific anticonvulsant drugs. In all, 40 of the 50 medical journals with the highest number of reports in Medline were hand-searched, together with nine specialist journals which were either not on that list or were not indexed.⁵⁶ The search process included volumes published between 1950 and 1990. Additional reports were identified from the reference list of the retrieved papers. Eligibility was determined by reading each report identified by the search. First authors were contacted for further information on the published report (method of randomisation, double-blinding, outcome measures and drop-outs) and to ask if they knew of any other studies which met our inclusion criteria, either performed by them or by other investigators.

Each report was scored for quality by four of the reviewers using a three-item scale (maximum 5, minimum 0);⁷ a joint 'consensus' score was then agreed for each report. Information on the pain condition and number of patients studied, anti-convulsant drug and dosing regimen, study design (placebo or active control), study duration and

follow-up, analgesic outcome measures and results, withdrawals and adverse effects was extracted from each report.

Dichotomous data was used to calculate odds ratio estimates with 95% CIs, using a fixed effects model,¹⁷¹ and an NNT⁶⁷ for effectiveness, for adverse effects and for drug-related study withdrawal. The index of effectiveness varied between reports. In some it was the numbers of patients who had improved, in others it was the numbers pain-free at the end of the study. No weighting was used between these different indices. The calculations were performed both for the individual reports and by combining single treatment or control arms. For the reports which presented mean data, with no significant difference between active and placebo, both treatments were assigned zero patients improved.

Results

A total of 37 reports were identified, all of which were published. Of these, 34 were identified from Medline and three from reference lists; 17 were excluded (*Table 23*). One was a duplicate publication.¹⁷²

TABLE 23 Reports excluded

Reference	Anti-convulsant	Condition	Reason for exclusion
Arieff & Wetzel, 1967 ¹⁹³	carbamazepine	neuralgias	not RCT
Farago, 1987 ³³⁷	carbamazepine	trigeminal neuralgia (TN)	not RCT
Fromm <i>et al</i> , 1984 ¹⁹⁴	carbamazepine	TN	not RCT
Goncikowska, 1984 ¹⁹⁵	carbamazepine	Horton's headache	not RCT
Hatta <i>et al</i> , 1992 ¹⁹⁶	phenytoin	suxamethonium-induced myalgia	suxamethonium-induced myalgia
Holmes <i>et al</i> , 1984 ¹⁹⁷	flunarizine	N/A	review, not RCT
Hopkins, 1991 ¹⁹⁸	carbamazepine	drug interaction	not RCT, drug interaction
Kienast & Boshes, 1968a	carbamazepine	TN	not RCT
Kienast & Boshes, 1968b ¹⁷²	carbamazepine	TN	not RCT, dual publication
Mathew & Ali, 1991 ¹⁹⁹	sodium valproate	chronic headache	not RCT
Naidu <i>et al</i> , 1991 ²⁰⁰	phenytoin	rheumatoid arthritis	not RCT
Rasmussen & Riishede, 1970 ²⁰¹	carbamazepine	facial pain	not RCT, single blind
Schaffler <i>et al</i> , 1991 ²⁰²	3-(4-hydroxypyridyl)-6-(2-chlorophenyl)-pyridazine	experimental pain in volunteers	experimental pain
Sharav <i>et al</i> , 1991 ²⁰³	carbamazepine	TN	case report not RCT
Shibasaki <i>et al</i> , 1973 ²⁰⁴	carbamazepine	Fabry's disease	case report not RCT
Westerholm, 1970 ²⁰⁵	carbamazepine	facial pain	not RCT
Young & Clarke, 1985 ²⁰⁶	clonazepam	diabetic neuropathy	not RCT

In all, 20 RCTs were considered eligible. Studies of four anticonvulsant drugs were identified, ten with carbamazepine, five with phenytoin, three with clonazepam and two with sodium valproate. Of the

20 studies, 17 related to chronic non-malignant pain, with one each related to cancer pain, postoperative pain and acute herpes zoster. Details of these eligible reports are in *Table 24* (placebo

TABLE 24 Reports included – trials with placebo control

Reference	Condition (number of patients)	Design, study duration, follow-up [#]	Outcome measures	Dosing regimen	Analgesic outcome results	Withdrawals and adverse effects	Quality score [†]
Carbamazepine							
Campbell <i>et al</i> , 1966 ¹⁷⁴	TN (77)	multicentre cross-over; 8 weeks (four 2-week periods, two each on carbamazepine & placebo); no follow-up.	pain severity, paroxysms, triggers	100 mg, q.d.s., to 200 mg t.d.s. (one centre) or 200 mg q.d.s. (two centres)	Mean % maximum possible pain intensity fell 58% with carbamazepine, 26% with placebo. Paroxysms & triggers also significantly reduced.	Seven withdrawals (1 rash, others logistic). 50% had one or more adverse effect on carbamazepine, 26% on placebo. Giddiness 30%, drowsiness 15%.	5
Killian & Fromm, 1968 ¹⁷⁵	TN (30), PHN (6), other chronic neuralgias (6). 36/42 studied double-blind (24/32 TN).	crossover; 10 days (two 5-day periods); open follow-up, range 2 weeks to 36 months.	pain relief	dose titration, 400 mg–1 g/day	19/27 TN 'complete or very good' response. Placebo responses 'minimal or absent in all cases'.	3/30 TN withdrawn (rash, leukopenia, abnormal liver function). Adverse effects in 23/36 studied double-blind; 17 giddiness, 16 drowsiness.	4
Nicol, 1969 ¹⁷⁶	64 facial pain recruited, 54 with TN. Results presented on 44 TN only, 'due to insufficient follow-up'.	partial crossover (successful first treatment period stayed on treatment); carbamazepine only – 20 placebo only – 7, placebo then carbamazepine – 17; follow-up, 46 months.	global rating (pain intensity & adverse effects)	dose titration, 100 mg–2.4 g/day	15/20 on carbamazepine ab initio had good or excellent response. 12/17 switched from placebo to carbamazepine and 6/7 on placebo also had good or excellent response.	2/37 carbamazepine withdrawn, 1 rash, 1 itch. 4/37 on carbamazepine died (of other causes). 10/37 drowsiness, 7/37 staggering gait.	3
Rull <i>et al</i> , 1969 ¹⁸⁰	diabetic neuropathy (30)	crossover; 6 weeks (three 2-week periods); no follow-up.	pain intensity	dose titration, 200 mg–600 mg/day	28/30 improved on carbamazepine vs 19/30 on placebo. 0/30 worse on carbamazepine, 11/30 worse on placebo.	Three withdrawals (2 carbamazepine adverse effects, 1 logistic). 16/30 somnolence, 12/30 dizziness.	4
Rompel & Bauermeister, 1970 ¹⁸³	migraine prophylaxis (48)	crossover; 12 weeks (two 6-week periods); no follow-up.	number of migraines, global rating	1 tablet t.d.s.*	38/45 improved on carbamazepine, 13/48 on placebo. 30 migraines in 45 on carbamazepine, 186 in 48 on placebo.	Three withdrawals, 1 adverse effects on carbamazepine, 2 logistic. 30/45 had adverse effects with carbamazepine (23 vertigo or dizziness), 11/48 on placebo.	4
Leijon & Boivie, 1989 ¹⁸⁷	central post-stroke pain (15)	crossover; 14 weeks (three 4-week periods with two 1-week washouts); no follow-up.	daily pain intensity, post-treatment global rating	stepped increase to final (day 18), 800 mg daily; amitriptyline; 75 mg daily at day 6.	Daily rating: amitriptyline significantly lower pain intensity than placebo on 3/3 weeks tested, and carbamazepine on 1/3 weeks tested. Global: 10/15 improved on amitriptyline, 5/14 on carbamazepine, 1/15 on placebo.	No withdrawals. 14/15 on amitriptyline and carbamazepine had adverse effects, 1/15 on placebo. Four patients on carbamazepine had dosage reduction.	4
[#] studies single centre unless stated; [†] quality score, maximum 5, minimum 0; * actual dose not specified.							

continued

TABLE 24 contd Reports included – trials with placebo control

Reference	Condition (number of patients)	Design, study duration, follow-up [#]	Outcome measures	Dosing regimen	Analgesic outcome results	Withdrawals and adverse effects	Quality score [†]
Phenytoin							
Greenbaum et al, 1973 ¹⁸⁶	irritable bowel (14)	crossover; 20 weeks (4 week run-in, 6 weeks treatment, 4 week washout, 6 weeks treatment); no follow-up.	bowel movements, pain episodes	fixed, 300 mg/day	No statistically significant differences between phenytoin and placebo; no quantitative data available.	Two withdrawals, neither drug-related. No report of adverse effects.	4
Chadda et al, 1978 ¹⁸¹	diabetic neuropathy (40)	crossover; 5 weeks (two 2-week periods, 1 week washout); no follow-up.	pain intensity; paraesthesiae	fixed, 300 mg/day	28/38 on phenytoin had at least moderate improvement, 10/38 on placebo.	Two withdrawals 'did not report back'. 4/38 giddiness on phenytoin.	2
Saudek et al, 1977 ¹⁸²	diabetic polyneuropathy (12)	crossover; 46 weeks (two x 23 weeks); no follow-up.	pain intensity	1 capsule t.d.s. titrated vs. plasma concentration*	No statistically significant differences between phenytoin and placebo. Mean pain score 7.2 mm for phenytoin plasma concentration < 5mg/l (placebo 8), and 19.1 vs. 20 for > 5 mg/l.	Two drug-related withdrawals on phenytoin, none on placebo. On phenytoin significant increase in plasma glucose and 4 x no. of reports of adverse effects (16 vs. 4).	2
Sodium valproate							
Martin et al, 1988 ¹⁹²	acute post-operative (39)	parallel group; (140 minutes).	pain intensity	15 mg/kg sodium valproate vs. placebo vs. 2 mg/kg ketoprofen, each i.v. over 20 minutes	No significant difference between valproate and placebo. Ketoprofen significant reduction in pain intensity (80–25 mm mean) compared with both valproate and placebo (80–60 mm).	None reported	2
Hering & Kuritzky, 1992 ¹⁸⁴	migraine prophylaxis (32)	crossover; 10 weeks (2-week run-in, two 8-week periods); no follow-up.	number of attacks, duration and pain intensity	fixed, 400 mg b.d.	Significant reduction in mean number of attacks (15 to 8), duration and pain intensity (24 to 15) on valproate. Valproate effective in 25/29.	Three withdrawals, 1 valproate adverse effects, 2 placebo. On valproate 2/29 dyspepsia, 2/29 nausea, 2/29 weariness.	3
Clonazepam							
Stensrud & Sjaastad, 1979 ¹⁸⁵	migraine prophylaxis (38)	crossover; 16 weeks (4-week run-in, three 4-week periods; placebo vs. high & low dose clonazepam); followed-up on 2 mg/day for 4 weeks.	headache days	fixed, 1 and 2 mg/day	Significant reduction in headache days between 1 (50%) and 2 mg (37%) and run-in period, but not for placebo (8%).	Four withdrawals, 3 lethargy on 1 mg clonazepam, unclear other. 23/38 drowsy and 10 dizzy on clonazepam.	3
Harkins et al, 1991 ¹⁸⁸	temporo-mandibular joint dysfunction (myofascial pain) (20)	parallel group; 60 days; open follow-up.	pain intensity on palpation and global pain	dose titration, 0.25–1 mg/day (mean 0.375 mg)	Significantly lower mean pain intensity and global pain at 30 days (vs. baseline) with clonazepam compared with placebo (about 40% change with clonazepam, 20% with placebo).	6/10 clonazepam withdrew; 1 at 1 week (headache), 5 at 30 days because pain improved. 7/10 placebo withdrew at 30 days because not improved. 3/10 clonazepam drowsy.	3
[#] studies single centre unless stated; [†] quality score, maximum 5, minimum 0; * actual dose not specified.							

controlled) and *Table 25* (active drug controls). The median quality score for the placebo-controlled studies was 3 (range, 2–5), and for the active drug controlled studies 2 (range, 1–4). Data were requested from 19 authors. Five replied, but only one (Leijon) was able to supply information relevant to this review.

Acute pain

Comparing carbamazepine with prednisolone in the management of acute herpes zoster, the patients treated with prednisolone reported less pain and faster skin healing (3.7 versus 5.3 weeks) than those treated with carbamazepine, 400 mg/day. In addition, 13/20 patients treated with carbamazepine still had pain at 2 months compared with 3/20 treated with prednisolone (*Table 25*).¹⁷³

Chronic pain

Trigeminal neuralgia

Three of the 12 placebo-controlled studies in chronic pain were in trigeminal neuralgia, all using carbamazepine (*Table 24* and *Figure 23*).^{174–176} Using dose titration to a maximum dose of 1 g/day, 19/27 patients had a complete or very good response compared with placebo on 5 days of treatment.¹⁷⁵ Again using dose titration and a crossover design, but to a maximum dose of 2.4 g/day, 15/20 patients randomised to initial carbamazepine had a good or excellent response compared with placebo on 14 days of treatment.¹⁷⁶ The extent to which the pain was relieved may be gauged from the third study.¹⁷⁴ Using doses in the range, 400–800 mg/day, for 2-week treatment periods, the mean fall in maximum pain intensity was 58% with carbamazepine compared with 26% on placebo. The NNT for

TABLE 25 Reports included – trials with active control

Reference	Condition (number of patients)	Design, study duration, comparator, follow-up [#]	Outcome measures	Dosing regimen	Analgesic outcome results	Withdrawals and adverse effects	Quality score*
Carbamazepine							
Keczkcs & Basheer, 1980 ¹⁷³	acute herpes zoster (40)	parallel group, 4 weeks vs. prednisolone; clinic follow-up till no pain (maximum > 1 year).	pain, skin healing, incidence of post-herpetic neuralgia (PHN) (> 2 months)	carbamazepine, 100 mg q.d.s. vs. prednisolone, 40 mg/day for 10 days then reduced to zero over next 3 weeks.	skin healing significantly faster with prednisolone. PHN in 13/20 carbamazepine patients, 3/20 prednisolone.	not reported.	1
Vilming et al, 1986 ¹⁷⁷	TN (12)	parallel group 3 weeks vs. tizanidine; no follow-up.	pain intensity, pain relief, global	carbamazepine titrated up to max of 3 x 300 mg/day, tizanidine to 3 x 6 mg/day.	4/6 carbamazepine patients rated treatment as very good, 1/5 tizanidine.	3 withdrawals on tizanidine, 1 not drug-related, 2 because of intolerable pain.	3
Lindstrom & Lindblom, 1987 ¹⁷⁸	TN (12)	crossover 4 weeks (two 2-week periods) vs. tocainide; no follow-up.	severity, frequency & duration of attacks = TN score	'maximum tolerated' dose of carbamazepine vs. tocainide 'about 20 mg/kg/day' in three divided doses.	tocainide and carbamazepine produced similar improvement compared with placebo – no significant difference between the active treatments.	tocainide 1/10 nausea, 1/10 paraesthesiae, 1/10 rash (withdrawn).	2
Lechin et al, 1989 ¹⁷⁹	TN (68)	multicentre (4), crossover 24 weeks (4-week placebo run-in then two 8-week periods with 4-week washout) vs. pimozone; open follow-up on pimozone.	TN symptom score	step titration carbamazepine, 300–1200 mg daily, and pimozone, 4–12 mg daily in two divided doses.	pimozone lowered symptom score by 78% from baseline compared with 50% on carbamazepine.	68 recruited, 59 randomised, 11 excluded from analysis, 10 protocol deviation, 1 did not return. 40/48 adverse effects on pimozone, 21/48 on carbamazepine.	4
[#] studies single centre unless stated; * quality score, maximum 5, minimum 0.							
							continued

TABLE 25 contd Reports included – trials with active control

Reference	Condition (number of patients)	Design, study duration, comparator, follow-up [#]	Outcome measures	Dosing regimen	Analgesic outcome results	Withdrawals and adverse effects	Quality score*
Phenytoin							
Richards et al, 1987 ¹⁸⁹	rheumatoid arthritis (60)	parallel group, single blind; 24 weeks; no follow-up.	pain and morning stiffness	step titration phenytoin, 100 mg/day, increased by 50 mg/week to effect or adverse effect. i.m. gold 10 mg test dose then 50 mg/week, then, if effective, 50 mg/2 weeks.	mean pain score improvement of 40/100 on gold, 12 on phenytoin.	6 withdrawals phenytoin, 2 no effect, 1 rash, 1 sleep difficulty, 1 lethargy. 1 unrelated death. 6 withdrawals on gold, 2 rash, 1 no effect, 3 proteinuria.	2
Yajnik et al, 1992 ¹⁹⁰	cancer pain (75)	parallel group; 4 weeks vs. buprenorphine plus combination of buprenorphine and phenytoin; no follow-up.	pain intensity, pain relief	phenytoin, 2 x 100 mg/day, buprenorphine, 2 x 0.2 mg sublingual/day, combination, 0.1 mg sublingual buprenorphine + phenytoin, 50 mg b.d.	good or moderate relief in 21/25 buprenorphine only, in 18/25 phenytoin only, and in 22/25 patients on combination therapy.	13/25 affected in buprenorphine only, 2/25 in phenytoin only, 5/25 in combination.	2
Combinations							
Gerson et al, 1977 ¹⁹¹	PHN (29)	parallel group; 8 weeks carbamazepine + clomipramine (16) vs. TENS (13); no follow-up.	pain intensity, activity, mood	carbamazepine, 150 mg–1 g /day, clomipramine 10–75 mg/day; TENS 15 min/week for 4 weeks.	mean improvement 43/100 mm on drug combination, 0.2 on TENS.	7 withdrawals on drug treatment, 4 (no effect) crossed (successfully) to TENS. 10 withdrawals on TENS (? cause) and 8 crossed to drug treatment, 3 successfully.	2

studies single centre unless stated; * quality score, maximum 5, minimum 0.

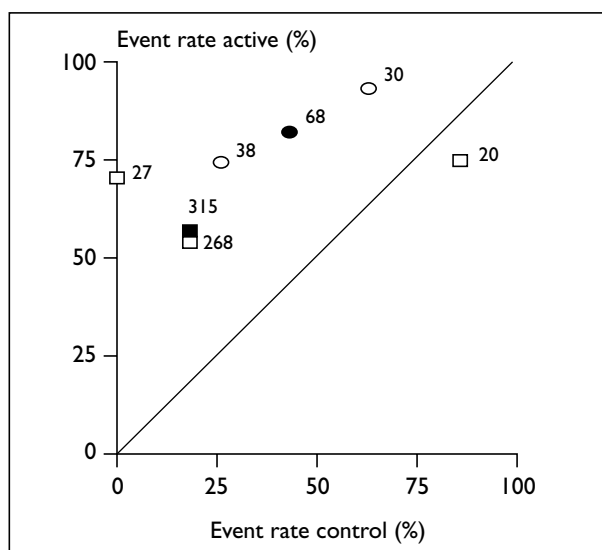


FIGURE 23 L'Abbé plot of TN and diabetic neuropathy trials with anticonvulsants, with numbers of patients shown for each trial (□, TN trials; ■, Combined TN trials; ○, Diabetic neuropathy trials; ●, Combined diabetic neuropathy trials)

effectiveness compared with placebo was 2.6, the NNH for adverse effects 3.4, and that for drug-related study withdrawal 24 (Table 26). The odds ratios for effectiveness of two of the three studies, and the overall ratio, showed carbamazepine to be better than placebo.

Three active control studies compared carbamazepine with tizanidine (alpha-2 adrenergic agonist),¹⁷⁷ tocainide (antiarrhythmic)¹⁷⁸ and pimozone (anti-psychotic) (Table 25).¹⁷⁹ Carbamazepine produced better results than tizanidine; there was no significant difference in the tocainide study; pimozone produced better results than carbamazepine.

Diabetic neuropathy

Two placebo-controlled studies in diabetic neuropathy (one with carbamazepine¹⁸⁰ and one with phenytoin¹⁸¹) found that after 2 weeks of treatment, 30–50% more patients improved on anticonvulsant

than on placebo. A third study using phenytoin for 23 weeks of treatment found no difference in mean pain intensity compared with placebo (Figure 23).¹⁸² The NNT for effectiveness compared with placebo was 3, the NNH for adverse effects 2.5, and for drug-related withdrawal from study 20 (Table 26). The odds ratio for overall effectiveness showed a significant effect for anticonvulsants compared with placebo.

There were no eligible active-controlled studies of diabetic neuropathy.

Migraine prophylaxis

Of three placebo-controlled studies of migraine prophylaxis, using three different anticonvulsants, two showed greater effect with the anticonvulsant than with placebo (Table 24). Treatment with carbamazepine, 3 tablets per day for 6 weeks, led to improvement in 38/45 patients compared with 13/48 on placebo.¹⁸³ Sodium valproate, 800 mg/day for 8 weeks, produced significant reduction in the number of migraines, in their duration and in pain intensity; valproate was effective in 25/29 patients.¹⁸⁴ The third study used clonazepam, 1 or 2 mg/day for

TABLE 26 Chronic pain reports with placebo control. NNT for effectiveness, NNH for adverse effects and drug-related withdrawals

Reference number	Effectiveness				Adverse effects				Drug-related study withdrawal			
	Improved on active drug	Improved on placebo	Odds ratio (95% CI)	NNT (95% CI)	On active drug	On placebo	Odds ratio (95% CI)	NNH (95% CI)	On active drug	On placebo	Odds ratio (95% CI)	NNH (95% CI)
TN												
174	144/268*	35/190*	4.4 (3–6.4)	2.8 (2.3–3.7)	38/77	20/77	2.7 (1.4–5.1)	4.3 (2.6–11.7)	1/77	0/77	7.4 (0.1–372.4)	77 (26.1–∞)
175	19/27	0/27	20.6 (6.8–62.3)	1.4 (1.14–1.88)	23/36	0/30	16 (5.8–43.9)	1.6 (1.3–2.1)	3/30	0/30	8 (0.8–79.3)	10 (4.8–∞)
176	15/20	6/7	0.5 (0.1–4)	–9.3 (4.7–∞)	10/37	0/7	4.5 (0.7–30.1)	3.7 (2.4–7.9)	2/37	0/7	3.4 (0.1–156.3)	18.5 (7.9–∞)
Combined	178/315	41/224	4.9 (3.4–6.9)	2.6 (2.2–3.3)	71/150	20/114	3.7 (2.2–6.2)	3.4 (2.5–5.2)	6/144	0/114	6.2 (1.2–31.7)	24 (13.5–110.8)
Diabetic neuropathy												
180	28/30	19/30	5.7 (1.7–19.2)	3.3 (2–9.4)	16/30	0/30	14.6 (4.7–45.5)	1.9 (1.4–2.8)	2/30	0/30	7.7 (0.5–126.6)	15 (6.4–∞)
181	28/38	10/38	6.5 (2.7–15.9)	2.1 (1.5–3.6)	4/38	0/38	8 (1.1–59.4)	9.5 (4.9–130)	0/38	0/38	n/a	n/a
182	no dichotomous data available				10/12	4/12	7.2 (1.5–35.3)	2 (1.2–6.3)	2/12	0/12	8 (0.5–136)	6 (2.6–∞)
Combined	56/68	29/68	5.4 (2.7–10.7)	2.5 (1.8–4)	30/80	4/80	6.9 (3.2–14.6)	3.1 (2.3–4.8)	4/80	0/80	7.7 (1.1–55.7)	20 (10.2–446)
Migraine prophylaxis												
183	38/45	13/48	9.9 (4.4–22.2)	1.7 (1.4–2.4)	30/45	11/48	5.8 (2.6–13.1)	2.3 (1.6–3.9)	1/48	0/48	7.4 (0.1–372.4)	48 (16.3–∞)
184	25/29	4/29	17.2 (6.2–47.7)	1.4 (1.1–1.8)	6/29	2/29	3.1 (0.7–13.7)	7.3 (3.2–∞)	1/32	2/32	0.5 (0.1–5)	(n/a)
185	no dichotomous data available				23/38	0/38	17 (6.4–44.8)	1.7 (1.3–2.2)	3/38	0/38	7.8 (0.8–77.4)	12.7 (6.1–∞)
Combined	63/74	17/77	6.1 (3.5–10.5)	1.6 (1.3–2)	59/112	13/115	6.7 (3.8–11.7)	2.4 (1.9–3.3)	5/118	2/118	2.4 (0.5–10.8)	39.3 (14.6–∞)
Other pain syndromes												
187	5/14	1/15	5.5 (0.9–32.3)	3.4 (1.7–105)	14/15	1/15	28.5 (7–116.5)	1.2 (1–1.5)	0/15	0/15	n/a	n/a
186	no dichotomous data available				/12	/12	n/a	n/a	0/12	0/12	n/a	n/a
188	no dichotomous data available				4/10	0/10	10.8 (1.3–91)	2.5 (1.4–10.4)	1/10	0/10	7.4 (0.1–372.4)	10 (3.5–∞)

* 77 patients assessed on multiple crossover. n/a not applicable or not available

60 days, and no significant difference was found between clonazepam and placebo.¹⁸⁵ The NNT for effectiveness compared with placebo was 2.4, the NNH for adverse effects 2.4, and for drug-related study withdrawal 39 (*Table 26*). The overall effectiveness odds ratio showed a significant effect for anticonvulsant compared with placebo.

There were no eligible active control studies of migraine prophylaxis.

Other pain syndromes

Placebo-controlled studies

Phenytoin, 300 mg/day for 6 weeks, had no effect in the one study of irritable bowel syndrome.¹⁸⁶ In central post-stroke pain, 4 weeks of carbamazepine at a final dose of 800 mg/day was judged to have improved 5/14 patients, compared with 10/15 patients on 75 mg of amitriptyline and 1/15 on placebo.¹⁸⁷ In a 60-day study of clonazepam (mean daily dose, 0.375 mg) in temporomandibular joint dysfunction, analysis at 30 days showed significantly lower pain intensity scores with the anticonvulsant compared with placebo.¹⁸⁸ Results from these studies are shown in *Table 24*.

Active-controlled studies

In rheumatoid arthritis, a 24-week comparison of phenytoin and intramuscular gold showed significantly better pain relief and morning stiffness in patients on gold at 24 weeks.¹⁸⁹ In cancer pain, phenytoin, 200 mg/day, was compared with buprenorphine alone and a combination of buprenorphine and phenytoin, 100 mg/day; all three regimens produced good or moderate relief in more than 60% of patients.¹⁹⁰ A comparison of a combination of carbamazepine and clomipramine with TENS produced improvement for 9/16 patients on drug treatment and 3/13 on TENS.¹⁹¹ Results from these studies are shown in *Table 25*.

Adverse effects and drug-related study withdrawal

In the placebo-controlled studies there were 16 drug-related patient withdrawals on anticonvulsant treatment compared with two on placebo (*Table 26*). Where adverse effects were reported, the incidence was between 25% and 50% within each study. Drowsiness, dizziness and gait disturbance were common problems.

Discussion

Process

Of the 37 reports found, 17 were excluded (15 because they were not RCTs). Many of the

remaining 20 had significant flaws; omission of drug dosage (1), lack of true blinding (1), inappropriate statistical conclusions (1), and omission of statistical testing (7). The quality scores of the placebo-controlled studies were higher than those of the active-controlled studies (*Tables 24* and *25*). Nine of the 20 papers were recent (published in the last decade) but standards of reporting have not improved. The quality of the reporting limited the ability to combine data, because many reports gave insufficient information and used a variety of different outcome measures, and several studies used variable dosing. Although the authors of the original reports were contacted by letter, not all replied and even those who did had no data available.

The NNT approach was used because most of the data were in dichotomous form, which lends itself to this analysis, and because NNTs are more readily clinically interpretable than, for instance, effect sizes. NNHs were calculated for adverse effects, minor and major, as well as the NNTs for effectiveness, because adverse effects are important for clinical decision-making. This approach may be useful in other reviews of long-established interventions. The older the report, the more likely it was to present simple binary data, such as improved versus not improved. More recent reports, which restricted data presentation to mean data for treatment and control, were not accessible to the NNT method.

Product

Acute pain

Anticonvulsants were ineffective in the one postoperative pain report,¹⁹² and in the acute herpes zoster report.¹⁷³ There is no logic in using anticonvulsant drugs to manage acute nociceptive pain when there are other (effective) remedies.

Chronic pain

The overall pattern, of NNT for effectiveness of about 2.5, NNH for adverse effects of about 3, and NNH for drug-related withdrawal from study between 20 and 40, was surprisingly similar for the three pain syndromes with more than one RCT (*Table 26*).

Medical students are often taught that a positive response to carbamazepine is 'diagnostic' for trigeminal neuralgia. If only one patient responds out of two treated, this statement needs to be qualified. One caveat is that the study populations may include patients who have had other interventions, such as nerve blocks, and the NNT for effectiveness may be more impressive in trigeminal

neuralgia treated with carbamazepine in the initial stages. The statement that “approximately 70% of patients will have significant pain relief”¹⁶² would seem to be about right.

Diabetic neuropathy is perceived as a model for other neuropathic pain syndromes, and results from diabetic neuropathy are often extrapolated to the other syndromes. The results with anticonvulsants reviewed here conflicted, with a negative result in the longest study (46 weeks) balanced by two positive studies. The NNT for effectiveness was the same as the NNH for adverse effects. The usual clinical decision is between antidepressant and anticonvulsant drugs as first-line treatment, and the evidence here does not support the use of anticonvulsants as first-line remedies. Direct evidence comparing antidepressant and anticonvulsant is available from the post-stroke pain study, where the NNT for effectiveness of amitriptyline was 1.7

compared with 3.4 for carbamazepine, with the same NNHs for adverse effects and study withdrawal.¹⁸⁷

The three placebo-controlled RCTs of anticonvulsant drugs in migraine prophylaxis showed anticonvulsants to be effective. Recent advances in migraine management may reduce the impact of these results.

This review shows that there is a need for high quality studies of the relative effectiveness of different anticonvulsant drugs in chronic pain syndromes, and for comparisons of antidepressant drugs with anticonvulsants. The usefulness of such primary studies would be increased greatly by improvements in the quality of reporting. Investigators presenting data as means for treatment and control should also consider the (simple) presentation of binary data.

Chapter 15

Antidepressants in neuropathic pain

Summary

The main outcomes from this review of RCTs on the effectiveness and safety of antidepressants in neuropathic pain were global judgements, pain relief or fall in pain intensity which approximated to more than 50% pain relief, and information about minor and major adverse effects. Dichotomous data for effectiveness and adverse effects were analysed using odds ratio and NNT methods.

A total of 21 placebo-controlled treatments in 17 RCTs were included, involving ten antidepressants. In six of 13 diabetic neuropathy studies the odds ratios showed significant benefit compared with placebo. The combined odds ratio was 3.6 (95% CI, 2.5–5.2), with an NNT for benefit of 3 (2.4–4). In two of three postherpetic neuralgia studies, the odds ratios showed significant benefit, and the combined odds ratio was 6.8 (3.5–14.3), with an NNT of 2.3 (1.7–3.3). In two atypical facial pain studies, the combined odds ratio for benefit was 4.1 (2.3–7.5), with an NNT of 2.8 (2–4.7). Only one of three central pain studies had analysable dichotomous data. The NNT point estimate was 1.7.

Comparisons of tricyclic antidepressants showed no significant differences between them; they were significantly more effective than benzodiazepines in the three comparisons available. Paroxetine and mianserin were less effective than imipramine.

For 11 of the 21 placebo-controlled treatments there was dichotomous information on minor adverse effects; combining across pain syndromes, the NNT for minor (noted in published report) adverse effects was 3.7 (2.9–5.2). Information on major (drug-related withdrawal from study) adverse effects was available from 19 reports; combining across pain syndromes, the NNT for major adverse effects was 22 (13.5–58).

Antidepressant drugs are effective in relieving neuropathic pain. With very similar results for anticonvulsant drugs, it is still unclear which class of drug should be first choice.

Introduction

For over 30 years antidepressants have been used to manage neuropathic pain but, in the UK, no antidepressant has a product licence for this indication.

Many of the studies of antidepressants in neuropathic pain are open case reports; interpreting the open studies is difficult because of the different drugs and doses used, and because of the simultaneous use of other drugs. The aim of this systematic review was to use the evidence from RCTs of antidepressants in neuropathic pain to address current clinical debates, which include which drug is best, whether selective serotonin reuptake inhibitors (SSRIs) have any advantages over tricyclic antidepressants, how to manage dose titration, whether benefit is due to analgesic effect rather than mood improvement, and whether the character of the pain is predictive of response, as well as to allow comparison with anticonvulsant drugs – the main therapeutic alternative.

In a previous review, Onghena and van Houdenhove²⁰⁷ looked at placebo-controlled studies in chronic non-malignant pain in general, rather than just neuropathic pain; they considered papers published up to 1990 and used effect size as the meta-analytic outcome. They concluded that “the average chronic pain patient who receives an antidepressant treatment is better off than 74% of the chronic pain patients who receive a placebo”. We wished to focus on neuropathic pain and, by using the NNT⁶⁷ as the meta-analytic outcome, to produce more precise clinical conclusions and allow comparison with the effect of anticonvulsants in neuropathic pain.⁷⁶

Methods

Reports of RCTs of antidepressants in chronic pain were sought. From these reports, the subset of trials of neuropathic pain (diabetic neuropathy, postherpetic neuralgia, atypical facial pain and central pain) were selected. Reports were included which were randomised comparisons of antidepressant with placebo, with another antidepressant, or with any other intervention. A number of different search strategies in both Medline (1966–94) and the Oxford

Pain Relief database (1950–92) were used to locate reports, using the individual drug names.³ Additional reports were identified from the reference lists of retrieved reports and from review articles. Lead authors of identified reports were contacted for more details and were asked if they knew of other reports.

Unpublished reports, abstracts and reviews, drugs withdrawn early in development and studies with fewer than ten patients per group were excluded. Two of the reviewers screened all reports to eliminate those which had no pain outcomes, were definitely not randomised, or were abstracts or reviews.

Each report which could possibly meet the inclusion criteria was read by each author independently and scored for inclusion and quality using a three-item scale.⁷ An included report could have a maximum score of 5 and a minimum score of 1. Information about the treatments and controls, type of condition studied, number of patients enrolled and analysed, study design, observation periods, outcome measures used for pain or global evaluation and their results, and minor (noted in published report) and major (drug-related study withdrawal) adverse effects was taken from each report, and agreed by all authors.

A clinically relevant outcome was defined as a measure equivalent to more than 50% of pain relieved. Dichotomous information was extracted for analysis. The effectiveness measures after the longest duration of treatment were used. A hierarchy of measures was used which approximated in this order:

- (i) patient global judgement (excellent/good)
- (ii) pain intensity (no pain/slight pain or > 50% decrease or from 'neuropathy' scale) or relief (good/excellent)

(iii) improved or marked improvement.

Analysis was undertaken separately for placebo- and active drug-controlled reports. Odds ratio estimates, the chance of the intervention being more effective than control, were used to answer the question, 'Does this intervention work compared with placebo?' and were calculated for each report, with 95% CIs, using a fixed effects model.²⁰⁸ The NNT, or NNH, was calculated with its 95% CIs for effectiveness and for minor and major adverse effects, both for the individual reports and by combining the data from the individual reports.⁶⁷ A statistically significant improvement over control was assumed when the lower limit of the 95% CI of the odds ratio was > 1. NNTs for effectiveness and adverse effects are reported with 95% CIs only when the odds ratio indicated a statistically significant improvement of the treatment over controls. Point estimates of the NNT without 95% CIs are reported when the odds ratio was not statistically significant. An infinity sign for the NNT in the tables indicates a negative or zero value. Calculations were performed using Excel v. 4.0 on a Power Macintosh 7100/80.

Results

The use of individual drug names (generic and brand) was required for maximum yield from searches of both Medline and the Oxford Pain Relief database. After excluding a number of reports as obvious reviews, experimental reports in humans or animals, or purely kinetic studies, a further ten reports were excluded at the final consensus meeting for the various reasons given in *Table 27*. Of the 18 reports on antidepressants that remained, several had multiple treatment arms, so that there were 21

TABLE 27 Papers excluded from analysis

Reference	Antidepressant	Reason for rejection
Davidoff <i>et al</i> , 1987 ²¹³	trazodone	Inadequate number of patients (9 per group)
Feinmann <i>et al</i> , 1983 ²¹⁴	dothiepin	Dual publication ²¹⁵
Feinmann <i>et al</i> , 1984 ²¹⁶	dothiepin	Dual publication ²¹⁵
Kvinesdal <i>et al</i> , 1983 ²¹⁷	imipramine	Dual publication ²¹⁸
Mendel <i>et al</i> , 1986 ²¹⁹	amitriptyline, fluphenazine	Inadequate number of patients (6)
Sharav <i>et al</i> , 1987 ²²⁰	amitriptyline	Mixed musculoskeletal and neurogenic
Sindrup <i>et al</i> , 1990 ²²¹	imipramine	Concentration-response study
Sindrup <i>et al</i> , 1992 ²²²	imipramine & paroxetine	No pain outcomes
Stockstill <i>et al</i> , 1989 ²²³	L-tryptophan	Inadequate number of patients (6–8 per group)
Young & Clarke, 1985 ²⁰⁶	imipramine	Inadequate number of patients (5)

placebo-controlled treatment arms, and 11 with active controls. Details of these studies and the quality scores are given in *Table 28A* (placebo-controlled) and *Table 28B* (active drug-controlled).

Results in placebo-controlled studies

The 21 eligible placebo-controlled treatment arms contained information on over 400 subjects

treated with ten different antidepressants and 373 subjects who received placebo (*Table 28*). Many of the studies had high scores on the quality scale.

The odds ratios and NNTs calculated for each treatment and for treatments combined across pain condition are shown in *Table 29*. In six of 13 reports in diabetic neuropathy, the odds ratios showed

TABLE 28 Antidepressant trial details

Study (Number of patients)	Design, study duration and follow-up	Dosing regimen	Outcome measures	Analgesic outcome results	Withdrawals and adverse effects	Quality score and comments
A. Placebo-controlled						
1 Diabetic neuropathy						
Gomez-Perez et al, 1985 ²²⁴ (24)	crossover, 2 x 30-day periods nortriptyline, in combination with fluphenazine	weeks 1–2: 10 mg nortriptyline + 0.5 mg fluphenazine 3 times/day. weeks 3–4: 20 mg nortriptyline + 1 mg fluphenazine 3 times/day or matching placebo	VAS pain intensity – decrease in pain as % of initial score	≥ 50% decrease in pain at 30 days: 16/18 nortriptyline + fluphenazine, 1/18 placebo (not significant at 15 days)	14/18 had adverse effects on nortriptyline + fluphenazine, 1/18 on placebo. No drug-related withdrawals	4
Kvinesdal et al, 1984 ²¹⁸ (15)	crossover, 2 x 5 week periods imipramine	week 1: imipramine, 50 mg/day. weeks 2–5: imipramine, 100 mg/day	1. 6-item neuropathy scale including pain – 3-point scale 2. global rating (patient & investigator)	1. neuropathy scale: no SD 2. patient global rating (improved): 7/12 imipramine, 0/12 placebo	12/15 had adverse effects on imipramine, 1/15 on placebo. 1/15 drug-related withdrawals on imipramine, 0/15 on placebo	4
Max et al, 1987 ²²⁵ (37)	crossover, 2 x 6 week periods amitriptyline	weeks 1–3: dose titration based on adverse effects, dose range: 25–150 mg. weeks 4–6: maintained on appropriate dose	1. verbal rating scale for pain intensity 2. patient global rating – 6-point scale 3. Hamilton Depression Scale	2. global rating (complete/virtually complete; a lot of pain relief): 15/29 amitriptyline, 1/29 placebo	28/29 had adverse effects on amitriptyline, 25/29 on placebo. 3/37 drug-related withdrawals on amitriptyline, 3/37 on placebo	4
Max et al, 1991 ²²⁶	crossover, 2 x 6 week periods desipramine vs benztrapine placebo	weeks 1–4: desipramine dose titration in 12.5–250 mg/day range; placebo (benztropine) 0.5–1.0 mg/day	1. CAT pain intensity – 4-point scale 2. patient global rating 3. Hamilton Depression Scale	1. desipramine significantly better than placebo 2. global rating (a lot/moderate pain relief): 11/20 desipramine, 2/20 placebo 3. depression improved on desipramine but not placebo	18/20 had adverse effects on desipramine, 17/20 on placebo. 2/24 drug-related withdrawals on desipramine, 1/24 on placebo	4
Max et al, 1992 ²¹¹ (46)	crossover, 2 x 6 week periods with 2-week washout fluoxetine vs benztrapine placebo	weeks 1–4: fluoxetine dose titration in 20–40 mg/day range, placebo (benztropine) 0.125–1.5 mg/day	1. CAT pain intensity – 13-point scale 2. patient global rating 3. Hamilton Depression Scale	1. no SD 2. global rating (complete/a lot/moderate pain relief): 22/46 fluoxetine, 19/46 placebo 3. depression improved on fluoxetine but not placebo	29/46 had adverse effects on fluoxetine, 31/46 on placebo. 3/46 drug-related withdrawals on fluoxetine, 2/46 on placebo	3

continued

TABLE 28 contd Antidepressant trial details

Study (Number of patients)	Design, study duration and follow-up	Dosing regimen	Outcome measures	Analgesic outcome results	Withdrawals and adverse effects	Quality score and comments
A. Placebo-controlled contd						
1 Diabetic neuropathy contd						
Sindrup <i>et al</i> , 1989 ²²⁷ (13)	crossover, 2 x 3 week periods imipramine	pre-study dose titration on 50 or 75 mg/day to estimate dose required to achieve target plasma concentration. Final doses range: 125–200 mg/day nocte	6-item neuropathy 5-point scale including pain	imipramine 8/9 scored 6 or less on neuropathy scale, placebo 7/9 8/9 scores lower on imipramine than on placebo	adverse effect score significantly higher on imipramine than on placebo. 1/9 drug-related withdrawals on imipramine, 2/9 on placebo, all for dizziness	4
Sindrup <i>et al</i> , 1990c ²²⁸ (26)	crossover, 3 x 2 week periods, 1–3 week washout clomipramine desipramine placebo	depending on metabolism: clomipramine 50–75 mg/day, desipramine 50–200 mg/day	6-item neuropathy 5-point scale including pain	clomipramine 14/18 scored 6 or less on neuropathy scale, desipramine 13/18, placebo 10/18	adverse effect score significantly higher on clomipramine and desipramine than placebo. 3/26 drug-related withdrawals on clomipramine, 3/26 on desipramine, 0/26 on placebo	4
Sindrup <i>et al</i> , 1990 ²²⁹ (26)	crossover, 3 x 2 week periods, 1–3 week washout paroxetine imipramine placebo paroxetine	40 mg/day imipramine adjusted to target plasma concentration (dose: 25–350 mg/day)	1. VAS pain intensity 2. 6-item neuropathy 5-point scale including pain 3. patient self-rating depression score	paroxetine 18/20 scored 6 or less on neuropathy scale, imipramine 18/19, placebo 14/20; depression: none of the treatments affected scores	adverse effects: only dry mouth significantly more common on imipramine than on paroxetine or placebo. 7/29 drug-related withdrawals imipramine, 0/29 on paroxetine, 0/26 on placebo	4
Sindrup <i>et al</i> , 1992b ²²⁹ (18)	crossover, 2 x 3 week periods, at least 1 week washout citalopram placebo citalopram	40 mg/day as a single dose at 20.00 hours	6-item neuropathy 5-point scale including pain	citalopram 13/15 scored 6 or less on neuropathy scale, placebo 8/15	adverse effect scores significantly higher on citalopram than placebo, median score 2 on citalopram, 0.04 on placebo. 2/18 drug-related withdrawals on citalopram, 0/18 on placebo	4
Sindrup <i>et al</i> , 1992c ²³⁰ (22)	crossover, 3 x 2 week periods, at least 1 week washout, double dummy mianserin imipramine placebo	imipramine adjusted to target plasma concentration (dose: 25–350 mg/day) mianserin 60 mg/day	6-item neuropathy 5-point scale including pain	imipramine 14/18 scored 6 or less on neuropathy scale, mianserin 11/18, placebo 11/18	adverse effect scores significantly higher on mianserin or imipramine than on placebo. 1/22 drug-related withdrawals on imipramine, 0/22 on mianserin, 0/22 on placebo	4
2 PHN						
Kishore-Kumar <i>et al</i> , 1990 ²³¹ (26)	crossover, 2 x 6 week periods desipramine placebo (benztropine)	weeks 1–4: dose titration dose range: desipramine 12.5–250 mg/day placebo (benztropine) 0.5–1.0 mg/day	1. CAT pain intensity – 13-point scale 2. global rating – 6-point scale	1. desipramine significantly better than placebo at weeks 3–6 2. global rating (moderate/a lot/ complete pain relief): 12/19 desipramine, 2/19 placebo; independent of mood	19/19 adverse effects on desipramine, 15/19 on placebo. 5/26 drug-related withdrawals on desipramine, 3/26 on placebo	4

continued

TABLE 28 contd Antidepressant trial details

Study (Number of patients)	Design, study duration and follow-up	Dosing regimen	Outcome measures	Analgesic outcome results	Withdrawals and adverse effects	Quality score and comments
A. Placebo-controlled contd						
2 PHN contd						
Max <i>et al</i> , 1988 ²³² (58)	crossover, 2 x 6 week periods amitriptyline lorazepam placebo	weeks 1–3: amitriptyline dose titration in range 12.5–150 mg/day lorazepam 0.5–6.0 mg/day	1. CAT pain intensity – 13-point scale 2. patient global rating – 6-point scale	1. amitriptyline significantly better than lorazepam and placebo 2. global rating (moderate/a lot/complete pain relief): 15/34 amitriptyline, 7/40 lorazepam, 5/25 placebo	30/34 adverse effects on amitriptyline, 18/25 on placebo. 5/34 drug-related withdrawals on amitriptyline, 3/25 on placebo	3
Watson <i>et al</i> , 1982 ²³³ (24)	crossover, 2 x at least 3 week periods, 1–2 week washout amitriptyline placebo	amitriptyline dose titration, starting dose 12.5 mg, dose range: 25–137.5 mg/day	1. VAS pain intensity 2. verbal rating of pain intensity 3. VAS depression 4. Beck Depression Inventory	1. amitriptyline significantly better than placebo 2. verbal pain scale (good/excellent): 16/24 amitriptyline 1/24 placebo 3. & 4. amitriptyline had no significant effect on depression	16/24 adverse effects on amitriptyline, 13/24 placebo. 1/24 drug-related withdrawals on amitriptyline, 0/24 placebo	2
3 Atypical facial pain						
Feinmann <i>et al</i> , 1984 ²¹⁵ (95)	parallel group, 3, 6 and 9 weeks, follow-up at 12 months dothiepin + nocturnal biteguard placebo + nocturnal biteguard dothiepin alone placebo alone	dothiepin dose titration in range 25–150 mg nocte. Mean daily dose 130 mg	1. frequency and severity of pain 2. reduction in analgesic use 3. Montgomery–Asberg depression rating scale	1. 34/48 dothiepin pain free (score 0/1 mild, occasional) at week 9, 21/45 placebo 2. reduction in analgesic use: 83% dothiepin 42% placebo 3. no SD in depression rating	adverse effects and drug-related withdrawals not stated	4 extra data available (letter from author)
Lascelles, 1966 ²³⁴ (40)	crossover, 2 x 1 month periods phenelzine placebo	phenelzine 3 x 15 mg/day	1. degree of pain 2. Hamilton Depression Scale	1. pain at 1 month (improvement/ marked improvement) 30/40 phenelzine, 9/40 placebo 2. depression score improved: 6/40 placebo, 12/40 phenelzine. Mood not independent of pain	adverse effects and withdrawals not stated	4
4 Central pain						
Leijon & Boivie, 1989 ¹⁸⁷ (post-stroke pain, N = 15)	crossover, 3 x 4 week periods, 1 week washout, double dummy amitriptyline carbamazepine placebo	amitriptyline day 1: 25 mg days 2–5: 50 mg days 6–28: 75 mg carbamazepine day 1: 200 mg days 2–5: 400 mg days 6–14: 600 mg days 15–17: 700 mg days 18–28: 800 mg	1. CAT pain intensity – 10-point scale 2. patient global rating – 5-point scale 3. Comprehensive Psycho-pathological Rating Scale	1. amitriptyline significantly better than placebo at weeks 2–4, carbamazepine significantly better than placebo at week 3 only 2. global rating 'improved' 10/15 amitriptyline, 5/14 carbamazepine, 1/15 placebo 3. depression: amitriptyline no significant effect	14/15 adverse effects on amitriptyline, 13/14 carbamazepine, 7/15 placebo. 0/15 drug-related withdrawals on all treatments	4

continued

TABLE 28 contd Antidepressant trial details

Study (Number of patients)	Design, study duration and follow-up	Dosing regimen	Outcome measures	Analgesic outcome results	Withdrawals and adverse effects	Quality score and comments
A. Placebo-controlled contd						
4 Central pain contd						
Panerai <i>et al</i> , 1990 ²³⁵ (central pain, N = 39)	crossover, 3 x 3 week periods clomipramine nortriptyline placebo	week 1: dose titration clomipramine 25–100 mg/day nortriptyline 25–100 mg/day	1. VAS pain intensity 2. investigator global rating – 4-point scale 3. Hamilton Depression Scale	1. both significantly better than placebo, clomipramine significantly better than nortriptyline 2. global rating: clomipramine > nortriptyline > placebo 3. depression: clomipramine > nortriptyline > placebo	23/24 adverse effects on clomipramine, 22/24 nortriptyline, 10/24 placebo. 0/39 drug-related withdrawals on clomipramine, 2/39 nortriptyline, 1/39 placebo	3 extra data available (letter from author)
B. Active-controlled						
Langohr, 1982 ²³⁶ (painful mono- and poly-neuropathies, N = 48)	crossover, 2 x 2 week periods, 1 week washout acetylsalicylic acid clomipramine	clomipramine, day 1: 50 mg/day, day 2: 100 mg/day, days 3–14: 150 mg/day; acetylsalicylic acid, day 1: 500 mg/day, day 2: 1000 mg/day, days 3–14: 1500 mg/day	1. effect of pain on well being, physical activity, sleep, walking – 5-point scale 2. global rating (patient & investigator)	doctor global rating 23/40 improved on clomipramine, 6/40 on placebo	37% adverse effects with clomipramine, 17% with acetylsalicylic acid. Withdrawals not stated	1
Turkington, 1980 ²³⁷ (diabetic neuropathy, N = 59)	parallel group, 3 months 2-year open follow-up	imipramine 100 mg nocte (20) amitriptyline 100 mg nocte (19) diazepam 5 mg tds (20)	1. painful legs (yes/no) 2. KDS depression scores	1. complete relief of leg pain in 20/20 on imipramine, 19/19 on amitriptyline and 0/20 on diazepam; 2. mean depression scores reduced significantly by imipramine or amitriptyline, not by diazepam	2/20 impotence or frigidity on imipramine, 1/19 diazepam; complete cure of all other symptoms by imipramine or amitriptyline, not by diazepam	2
Watson <i>et al</i> , 1992 ²³⁸ (PHN, N = 35)	crossover, 2 x 5 week periods, 2 week washout amitriptyline maprotiline	weeks 1–3: dose titration, starting dose 12.5 mg; weeks 4–5: stable dose	1. VAS pain intensity 2. CAT pain intensity – 4-point scale 3. % pain relief 4. Beck Depression Inventory 5. clinical effectiveness – 4-point scale	1. amitriptyline significantly better than maprotiline 2. mild/no pain: 15/32 amitriptyline 12/32 maprotiline 3. no SD in % pain relief 4. depression: no significant effect of amitriptyline or maprotiline 5. amitriptyline clinically more effective	20/35 adverse effects on amitriptyline, 28/35 maprotiline. 3/35 drug-related withdrawals on amitriptyline, 3/35 maprotiline	4

significant benefit compared with placebo. The combined odds ratio for all 13 reports was 3.6 (95% CI, 2.5–5.2), with a NNT of 3 (2.4–4). The 13 studies used nine different antidepressants. Some drugs were more effective than others. The NNT for the four combined imipramine hydrochloride reports was 3.7 (2.3–9.5), for the two desipramine reports 3.2 (1.9–9.7) and for the eight combined tricyclic reports 3.2 (2.3–4.8). In contrast, the point estimate NNT for paroxetine was 5, for fluoxetine

15.3, and for mianserin there was no difference from placebo. An important feature was the very considerable variation in the improvement seen on placebo (see *Table 29* and *Figure 24*). This variation ranged from 0% to 75% within the 13 diabetic neuropathy studies.

In two of three studies in postherpetic neuralgia the odds ratios showed significant benefit, and the combined odds ratio was 6.8 (3.4–14.3), with an

TABLE 29 NNT for benefit, minor and major harm in placebo-controlled trials

Trial	Benefit				Minor harm			Major harm		
	Active drug	Improved on active (on control)	Odds ratio (95% CI)	NNT (95% CI)	Adverse effects on active (on control)	Odds ratio (95% CI)	NNT (95% CI)	Withdrawal on active (on control)	Odds ratio (95% CI)	NNT (95% CI)
Diabetic neuropathy										
Gomez-Perez et al, 1985 ²²⁴	nortriptyline fluphenazine combination	16/18 (1/18)	25.8 (7.1–93.8)	1.2 (1–1.5)	14/18 (1/18)	18 (4.9–66.4)	1.4 (1.1–2)	0/18 (0/18)	∞	∞
Kvinesdal et al, 1984 ²¹⁸	imipramine	7/12 (0/12)	15 (2.7–83.9)	1.7 (1.2–3.3)	12/15 (1/15)	17.9 (4.3–74.2)	1.4 (1–2)	1/15 (0/15)	7.4 (0.1–372.4)	15 (5.2–∞)
Max et al, 1987 ²²⁵	amitriptyline	15/29 (1/29)	10.8 (3.4–33.7)	2.1 (1.5–3.5)	28/29 (25/29)	3.6 (0.6–22.4)	9.7 (4.1–∞)	3/37 (3/37)	1 (0.2–5.2)	∞
Max et al, 1991 ²²⁶	desipramine	11/20 (2/20)	7.4 (2–27.3)	2.2 (1.4–5.1)	18/20 (17/20)	1.6 (0.2–9.9)	20 (3.9–∞)	2/24 (1/24)	2 (0.2–20.3)	24 (5.6–∞)
Max et al, 1992 ²¹¹	fluoxetine	22/46 (19/46)	1.3 (0.6–2.9)	15.3 (3.7–∞)	29/46 (31/46)	0.8 (0.4–1.9)	–	3/46 (2/46)	1.5 (0.3–9.1)	46 (8.8–∞)
Sindrup et al, 1989 ²²⁷	imipramine	8/9 (7/9)	2.1 (0.2–23.7)	9 (2.2–∞)	no dichotomous data available			1/9 (2/9)	∞	∞
Sindrup et al, 1990c ²²⁸	clomipramine	14/18 (10/18)	2.6 (0.7–10.4)	4.5 (1.9–∞)	no dichotomous data available			3/26 (0/26)	8 (0.8–80.7)	8.7 (4.2–∞)
Sindrup et al, 1990c ²²⁸	desipramine	13/18 (10/18)	2 (0.5–7.7)	6 (2.1–∞)	no dichotomous data available			3/26 (0/26)	8 (0.8–80.7)	8.7 (4.2–∞)
Sindrup et al, 1990 ²⁰⁹	imipramine	18/19 (14/20)	5.1 (1–25.8)	4 (2.1–43.8)	no dichotomous data available			7/29 (0/29)	9.3 (2–44.8)	–
Sindrup et al, 1990 ²⁰⁹	paroxetine	18/20 (14/20)	3.4 (0.7–15.6)	5 (2.3–∞)	no dichotomous data available			0/29 (0/29)	∞	–
Sindrup et al, 1992b ²²⁹	citalopram	13/15 (8/15)	4.6 (1–21.5)	3 (1.6–35.9)	no dichotomous data available			2/18 (0/18)	7.8 (0.5–130.5)	9 (3.9–∞)
Sindrup et al, 1992c ²³⁰	mianserin	11/18 (11/18)	1 (0.3–3.7)	–	no dichotomous data available			0/22 (0/22)	∞	∞
Sindrup et al, 1992c ²³⁰	imipramine	14/18 (14/18)	2.1 (0.5–8.7)	6 (2.2–∞)	no dichotomous data available			1/22 (0/22)	∞	22 (7.5–∞)
Combined *		180/26 (72/205)	3.6 (2.5–5.2)	3 (2.4–4)	101/128 (27/62)	12.6 (5.4–30)	2.8 (2–4.7)	26/321 (5/174)	3.3 (1.4–7.8)	19.1 (11–74.4)
PHN										
Kishore-Kumar et al, 1990 ²³¹	desipramine	12/19 (2/19)	9 (2.5–33.2)	1.9 (1.3–3.7)	19/19 (15/19)	8.8 (1.1–68.1)	4.8 (2.5–36.8)	5/26 (3/26)	1.8 (0.4–7.9)	13 (3.7–∞)
Max et al, 1988 ²³²	amitriptyline	15/34 (5/25)	2.9 (1–8.5)	4.1 (2.1–82.1)	30/34 (18/25)	2.9 (0.8–10.7)	6.2 (2.7–∞)	5/34 (3/25)	1.3 (0.3–5.6)	37 (5–∞)
Watson et al, 1982 ²³³	amitriptyline	16/24 (1/24)	14.5 (4.5–46.8)	1.6 (1.2–2.4)	16/24 (13/24)	1.7 (0.5–5.2)	8 (2.5–∞)	1/24 (0/24)	7.4 (0.1–372.4)	24 (8.2–∞)
Combined		43/77 (8/68)	6.8 (3.4–14.3)	2.3 (1.7–3.3)	65/77 (46/68)	2.5 (1.2–5.4)	6 (3.3–33.2)	11/84 (6/75)	1.7 (0.6–4.6)	19.6 (6.9–∞)
Atypical facial pain										
Feinmann et al, 1984 ²¹⁵	dothiepin	34/48 (21/45)	2.7 (1.2–6.1)	4.1 (2.3–21.1)	no dichotomous data available			not stated		
Lascelles, 1966 ²³⁴	phenelzine	28/40 (9/40)	6.6 (2.8–15.8)	2.1 (1.5–3.5)	no dichotomous data available			not stated		
Combined		62/88 (30/85)	4.1 (2.3–7.5)	2.8 (2–4.7)						
<i>continued</i>										
* Placebo numbers counted once only for Sindrup, 1990a, ²⁰⁹ 1990b, ²²¹ 1992b, ²²⁹ and Panerai, 1990, ²³⁵ Max, 1991, ²²⁶ 1992 ²¹⁰ omitted from adverse effects (active placebo); effectiveness assessed for longest recorded period; ∞, zero or negative value.										

TABLE 29 contd NNT for benefit, minor and major harm in placebo-controlled trials

Trial	Benefit			Minor harm			Major harm			
	Active drug	Improved on active (on control)	Odds ratio (95% CI)	NNT (95% CI)	Adverse effects on active (on control)	Odds ratio (95% CI)	NNT (95% CI)	Withdrawal on active (on control)	Odds ratio (95% CI)	NNT (95% CI)
Central pain										
Leijon & Boivie, amitriptyline 1989 ¹⁸⁷		10/15 (1/15)	12.2 (2.8–52.3)	1.7 (1.1–3)	14/15 (7/15)	8.6 (1.8–39.8)	2.1 (1.3–5.4)	0/15 (0/15)	∞	∞
Panerai et al, 1990 ²³⁵	clomipramine	no dichotomous data available			23/24 (10/24)	10/24	11.8 (3.5–39.5)	0/39 (1/39)	0.1 (0–6.8)	∞
Panerai et al, 1990 ²³⁵	nortriptyline	no dichotomous data available			22/24 (10/24)	10/24	9.1 (2.8–29.7)	2/39 (1/39)	2 (0.2–19.6)	39 (9–∞)
Combined		59/63 (17/39)	10.5 (4–27)	2 (1.5–3)	2/93 (1/54)	1 (0.1–7.2)	334.8 (20.2–∞)			

* Placebo numbers counted once only for Sindrup, 1990a,²⁰⁹ 1990b,²²¹ 1992b,²²⁹ and Panerai, 1990,²³⁵ Max, 1991,²²⁶ 1992²¹⁰ omitted from adverse effects (active placebo); effectiveness assessed for longest recorded period; ∞, zero or negative value.

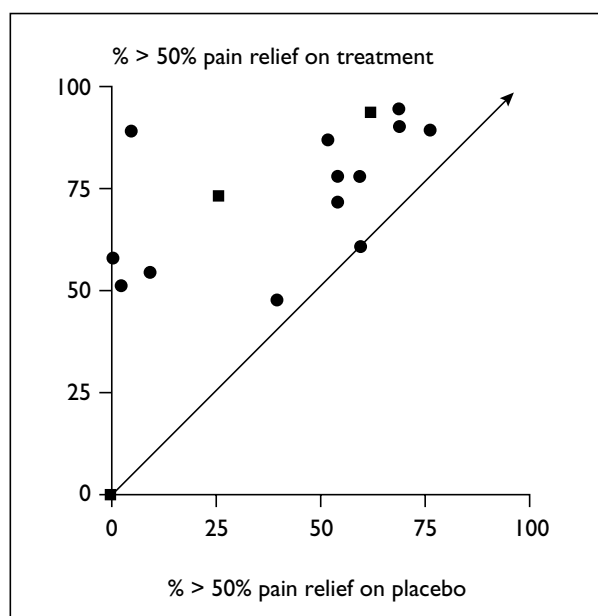


FIGURE 24 L'Abbé plot for trials⁷⁶ of antidepressants in chronic pain from diabetic neuropathy (● antidepressants, ■ anticonvulsants)

NNT of 2.3 (1.7–3.3). In both atypical facial pain studies, the odds ratios showed significant benefit, and the combined odds ratio was 4.1 (2.3–7.5), with an NNT of 2.8 (2–4.7). Only one of three studies in central pain had analysable dichotomous data. The NNT point estimate for benefit was 1.7 in 15 treated patients. The overall NNT for tricyclic antidepressants, combining the data from 13 studies across pain condition, was 2.9 (2.4–3.7).

Adverse effects

Dichotomous information on minor adverse effects was available from 11 of the 21 placebo-controlled reports (Table 29). Again, there was considerable variation across the placebo treatments, with minor adverse effect incidence varying from 6% to 86%.

Combining across pain syndromes (11 reports), the NNT for minor adverse effects was 3.7 (2.9–5.2).

Dichotomous information on major adverse effects was available from 19 of the 21 placebo-controlled studies (Table 29); drug-related withdrawals occurred in 41/498 (8%) patients treated with an active drug compared with 14/303 (5%) patients treated with placebo. Combining across pain syndromes (19 reports), the NNT for major adverse effects was 22 (13.5–58). This combined figure conceals a lower incidence in the studies for SSRIs (fluoxetine and paroxetine) than for tricyclics.

Results in active drug-controlled studies

In the three reports which compared tricyclic antidepressants with benzodiazepines, tricyclics were significantly more effective. The two studies with dichotomous data on comparisons of different tricyclics did not show any significant difference. The difference between imipramine and paroxetine was not statistically significant and had an NNT point estimate of 21. The comparison between imipramine and mianserin did not show a significant odds ratio but had an NNT point estimate of 6. There was no evidence of different minor or major adverse effect incidence for the various drugs. (See Tables 28 and 30 for fuller details.)

Discussion

Antidepressants clearly have an analgesic effect when compared with placebo in neuropathic pain. This effect was apparent for several different pain syndromes, and was of a similar magnitude in the different syndromes, despite the presumed differences in the underlying pain mechanisms. Compared with placebo, of 100 patients with

TABLE 30 NNT for benefit in active-controlled trials

Trial	Active drug	Improved on active drug	Improved on control	Odds ratio (95% CI)	NNT (95% CI)
Diabetic neuropathy					
Langohr <i>et al</i> , 1982 ²³⁶	clomipramine/acetysalicylic acid	23/40	6/40	6.1 (2.5–15.2)	2.4 (1.6–4.2)
Max <i>et al</i> , 1992 ²¹¹	amitriptyline/desipramine	28/38	23/38	1.8 (0.7–4.7)	7.6 (2.9–∞)
Sindrup <i>et al</i> , 1990 ^{c228}	clomipramine/desipramine	14/18	13/18	1.3 (0.3–5.9)	18 (3–∞)
Sindrup <i>et al</i> , 1990 ²⁰⁹	imipramine/paroxetine	18/19	18/20	1.9 (0.2–19.6)	21.1 (4.7–∞)
Sindrup <i>et al</i> , 1992 ²³⁰	imipramine/mianserin	14/18	11/18	2.1 (0.5–8.7)	6 (2.2–∞)
Turkington, 1980 ²³⁷	imipramine/diazepam	20/20	0/20	∞	∞
Turkington, 1980 ²³⁷	amitriptyline/diazepam	19/19	0/20	∞	∞
Combined *		136/172	71/174	6.1 (3.9–9.6)	2.6 (2.1–3.5)
PHN					
Max <i>et al</i> , 1988 ²³²	amitriptyline/lorazepam	15/34	7/40	3.5 (1.3–9.5)	3.8 (2.1–16.2)
Watson <i>et al</i> , 1992 ²³⁸	amitriptyline/maprotiline	15/32	12/32	1.5 (0.5–3.9)	10.7 (3–∞)
Combined		30/66	19/72	2.3 (1.1–4.5)	5.3 (2.9–30)
Central pain					
Leijon & Boivie, 1989 ¹⁸⁷	amitriptyline/carbamazepine	10/15	5/14	3.3 (0.8–13.9)	3.2 (1.5–∞)
Panerai <i>et al</i> , 1990 ²³⁵	clomipramine/nortriptyline	no dichotomous data available			N/A

* Effectiveness assessed for longest recorded period; N/A, not applicable; ∞, zero or negative value.

neuropathic pain who are given antidepressants, 30 will obtain more than 50% pain relief, 30 will have minor adverse reactions, and four will have to stop treatment because of major adverse effects.

Within this overall pattern, SSRIs were less effective in two studies than tricyclic antidepressants. There was insufficient data to say whether SSRIs caused fewer minor adverse effects, but the rate of major adverse reactions was half that seen with the tricyclics (*Table 29*).

Antidepressant studies in chronic pain are not easy to perform, and one issue is whether to study a fixed dose, avoiding the difficult problem of titrating to an effective dose before embarking on a trial of that effective dose against control, or to do the pre-trial titration. Some studies in this systematic review titrated pre-trial, others used a fixed dose (*Table 28*). No obvious difference in outcome was apparent for the two approaches. In clinical practice, titration to benefit and minimal adverse effect can be performed rapidly with tricyclics, with response evident within 5 days and perhaps less.^{209,210}

While controversy has continued as to whether the analgesic effect is separable from the effect of

the antidepressants on mood, many of these studies showed analgesic benefit without significant change in mood measurements. Another point of debate, that the character of the pain is predictive of response, may also be resolving. The adage that burning pain should be managed with antidepressant and shooting pain with anticonvulsant is not supported. If benefit was found, it occurred independent of pain character.²¹¹

The variation in the response to placebo groups in the different trials (*Figure 24*) is intriguing. Plotting for each trial the response to treatment (y axis) against response to placebo (x axis) shows that a higher proportion of patients achieved more than 50% relief on active treatment (antidepressant or anticonvulsant⁷⁶) than on placebo so that, for most trials, the points are plotted in the upper left section of the figure. Both anticonvulsants and antidepressants were effective in diabetic neuropathy, 50–85% of patients achieving more than 50% pain relief. In response to placebo, the proportion of patients with greater than 50% pain relief varied from 0% to 75%. Overall, the variation in response to placebo was greater than the response to treatment. We have no simple explanation as to why, within this set of trials, on a supposedly homogeneous population of patients

all with diabetic neuropathy, such variation should occur. The Hawthorne effect, a change in patient behaviour due to participation in a trial,²¹² may well apply within a particular trial, elevating patient response to placebo because of the extra attention, but then there is a variation in the extent of the Hawthorne effect within this set of trials – some had greater responses than others.

The aim of systematic reviews should be to guide clinicians to the most effective intervention in a particular condition. In *Figure 24* the points plotted for diabetic neuropathy trials of anticonvulsants⁷⁶ and for antidepressants (this review) show similar scatter, suggesting that from the available trials there is no measurable difference in the analgesic benefit of the two classes of drug in neuropathic pain. The combined NNT for benefit for antidepressants in diabetic neuropathy was 3 (95% CI, 2.4–4), and for anticonvulsants 2.5 (95% CI, 1.8–4).⁷⁶ There were no substantial difference in the adverse effects, minor or major, in the diabetic neuropathy patients either. The NNT for minor adverse effects of antidepressants was 2.8 (2–4.7),

for anticonvulsants 3.1 (2.3–4.8), for severe adverse effects with antidepressants 19 (11–74.4), and with anticonvulsants 20 (10–446).

Many clinicians prescribe antidepressants rather than anticonvulsants as first-line therapy in neuropathic pain, either because of perceived greater chance of benefit or lower chance of adverse effects. The only randomised comparison of antidepressant with anticonvulsant showed greater benefit at lower risk with antidepressant.¹⁸⁷ This stratagem is not supported by the systematic reviews, which indicate little to choose between antidepressant and anticonvulsant. This then is a straightforward research agenda. We need to determine the relative risk and benefit of the best and most appropriate anticonvulsant and the best and most appropriate antidepressant, and then compare them directly in neuropathic pain. A further agenda would be to see if the combination of anticonvulsant and antidepressant performed better than either component alone. Another enterprise would be to determine which aspects of care increase the proportion of patients taking placebo who achieve better than half relief.

Chapter 16

Systemic local anaesthetic type drugs in chronic pain

Summary

In order to review the effectiveness of systemically-administered local anaesthetic type drugs in chronic pain, RCTs with pain outcomes were sought; 21 RCTs were found. Three duplicate publications were excluded, as was one study where randomisation was not stated. Of the remaining 17 studies, ten used intravenous lignocaine, two intranasal lignocaine, four oral mexiletine and one oral tocainide (a total of 450 patients were included).

In pain due to nerve injury, intravenous lignocaine was effective in all four studies, showing either significant pain relief over placebo or a positive dose response. Oral mexiletine was shown to be more efficacious than placebo in all three studies in pain due to nerve damage but lacked effect in central pain due to spinal cord injury. Allodynia in pain due to peripheral nerve injury was relieved by intravenous lignocaine, as was dysaesthesia due to diabetic neuropathy, in which oral mexiletine was also effective. Intravenous lignocaine showed some efficacy in fibromyalgia but no effect was demonstrated in any of the three studies in cancer pain. It was inconsistent in migraine.

The best documented effective dose of intravenous lignocaine was 5 mg/kg, which was well tolerated when infused over 30 minutes. Mexiletine (225–750 mg) caused minor adverse effects that were dose-related. Tocainide should not be used because of toxicity.

Local anaesthetic type drugs are effective in pain due to nerve damage but there is little or no evidence to support their use in migraine or cancer-related pain.

Introduction

After peripheral nerve injury neuromas can be formed. Both the neuroma and the dorsal root ganglion display spontaneous activity and increased sensitivity to chemical and mechanical stimuli.^{239–242} In experimental models of nerve injury, systemic sodium channel blockers like lignocaine and

mexiletine silence spontaneous activity of neuroma and dorsal root ganglion, and reduce their mechanosensitivity at concentrations that do not block nerve conduction.^{243–245} Low doses of lignocaine may block glutamate-evoked activity in the dorsal horn of the spinal cord.²⁴⁶

Lignocaine and related local anaesthetic type drugs which block sodium channels have therefore been used to relieve clinical pain, as a last resort in cancer-related pain and in other conditions when more traditional treatments have failed.

Methods

RCTs of local anaesthetic type drugs were sought. Reports were included if they were randomised comparisons of local anaesthetic type drugs with a placebo or/and an active control. A number of different search strategies in Embase, Medline Knowledge Server, Silver Platter (1966–September 96) and the Oxford Pain Relief database (1950–94)⁶ were used without language restriction. Search terms (free text) included: (mexiletine/mexitil/flecainide/tambacor/lignocaine/lidocaine/ xylocaine/ xylocard/tocainide/procainide/pronestyl/ encainide) and (pain/painful/analgesic/analgesia).

Additional reports were identified from the reference lists of retrieved reports and from review articles. Unpublished reports, abstracts, reviews or reports of experimental pain were not considered. Authors were not contacted.

Each report was read by each of the reviewers independently to address methods of randomisation and blinding, and description of withdrawals.⁷ Reviewers then reached a joint consensus. The minimum quality score of an included RCT was 1, the maximum 5. There was a *pre-hoc* agreement that trials without randomisation or with an inadequate randomisation method (without concealment of treatment allocation) would be excluded from further analysis.

Information about the treatments and controls, characteristics of the pain condition, number of

patients enrolled and analysed, study design, observation periods, outcome measures used for pain intensity, pain relief and consumption of supplementary analgesics and adverse effects were taken from each report. Quantitative analysis was attempted. An NNT was calculated⁶⁷ when possible and a L'Abbé plot was constructed to analyse the degree of pain relief in different pain conditions.⁷⁴ Adverse effects were considered major if they necessitated discontinuation of the treatment or lowering of the dose.

Results

A total of 21 RCTs of systemic local anaesthetic type drugs in pain relief were found. Three reports were published twice, leading to three exclusions,²⁴⁷⁻²⁴⁹ and one report²⁵⁰ combined one randomised with one non-randomised study. The remaining 17 studies were in neuropathic pain (nine studies), fibromyalgia (one study), facial pain (one study), cancer pain (three studies) and acute migraine (three studies) (*Table 31*).

TABLE 31 Details of RCTs

Study	Condition; number of patients analysed per group	Design; duration; follow-up	Pain relief	Adverse effects	Quality score
Peripheral nerve injury					
Chabal <i>et al</i> , 1992 ²⁴³	Peripheral nervous system injury/disease, n = 11 (men only) duration of symptoms 1–30 year. 5/11 on tricyclic antidepressants, 1/11 on NSAID all had failed to receive adequate. Pain relief after conventional methods	Crossover, double blind, titration to maximum tolerable dose, then stable dose for 4 weeks oral mexiletine vs. placebo	Mexiletine, 450 mg, significant reduction from baseline, 750 mg SD from placebo, indication of some dose response	Mexiletine: 2/11 mild nausea; placebo: none reported	3
Galer <i>et al</i> , 1996 ²⁵¹	Chronic neuropathic pain, n = 9, (allodynia 5/9): polyneuropathy (n = 6), local nerve injury (n = 2), arachnoiditis (n = 1)	Crossover, double blind, ≥ 1 week between treatments lignocaine, 2 mg/kg vs. lignocaine, 5 mg/kg	VAS pain intensity: significant reduction ($p < 0.05$) from baseline with both doses, positive dose-response in VAS pain relief but not in pain intensity	1/9 withdrew on both doses: heavy feelings and general weakness (conc. 1.1 mg/l), extreme light-headedness & tinnitus (conc. 2.1 mg/l)	4
Wallace <i>et al</i> , 1996 ²⁵²	Peripheral nerve injury, n = 11	Crossover, double blind, 1 week between treatments lignocaine i.v. vs. NSAID i.v.	Positive concentration dependent reduction in VAS pain intensity (≥ 1.5 mg/l)	6/11 light-headedness (1.5 ± 0.6 mg/l), 1/11 nausea (2.3 mg/l), 2/11 discontinuation of the infusion; placebo: 1/11	2
Diabetic neuropathy					
Dejgard <i>et al</i> , 1988 ²⁵³	Diabetic neuropathy (duration > 6 m) n = 16	Crossover, double blind, 4 weeks treatment, 1 week washout post-operative mexiletine vs. placebo. Ibuprofen/acetaminophen/ non-pain medication continued, other previous pain medication discontinued	Mexiletine: sign pain relief, dose response not looked at, no significant correlation between plasma concentration and clinical effect	Mexiletine: 3/16 mild adverse effects: nausea, hiccup, tremor (one lowered the dose); placebo: none reported	3

continued

TABLE 31 contd Details of RCTs

Report	Condition; number of patients analysed per group	Design; duration; follow-up	Pain relief	Adverse effects	Quality score
Diabetic neuropathy contd					
Kastrup <i>et al</i> , 1987 ²⁵⁴	Diabetic neuropathy, n = 15	Crossover, double blind, 2 x 5 weeks, 2 week washout before study, no other analgesics allowed lignocaine i.v. vs. placebo i.v.	< 15 mm reduction in VAS pain intensity: lignocaine: 11/15 vs. placebo: 4/15	No adverse effects noted	2
Stracke <i>et al</i> , 1992 ²⁵⁵	Diabetic neuropathy (lasting > 4 months and < 5 years with pain score \geq 25% on VAS pain intensity during washout week). Mexiletine: n = 47; placebo: n = 48	Parallel group, 1 week washout, 6 weeks treatment, paracetamol allowed; post operative mexiletine vs. placebo. Multicentre (7 centres)	No overall pain relief but significant effect in subgroups (burning, stabbing and heat sensations)	675 mg/day: mexiletine: 11/47, placebo: 6/48, 450 mg/day: more adverse effects with placebo	2
PHN					
Rowbotham <i>et al</i> , 1991 ²⁵⁶	PHN (duration > 3 months) n = 19, allodynia: 16/19	Crossover, double blind, 3 x 3 hour, at least 48 hours between treatments lignocaine i.v. vs. morphine i.v. vs. placebo i.v.	VAS pain intensity lignocaine and morphine significantly better than placebo ($p < 0.05$), VAS pain relief: morphine significantly better than placebo ($p = 0.01$); lignocaine vs. placebo ($p = 0.06$)	Lignocaine: 1/19 nausea and light-headedness at 180 mg, conc. 1.33 mg/l; morphine: 7/19 nausea	3
Dysaesthetic spinal cord injury					
Chiou-Tan <i>et al</i> , 1996 ²⁵⁷	Dysaesthetic spinal cord injury, n = 11	Crossover, double blind, 4 weeks treatment, 1 week washout. Oral mexiletine vs. placebo ibuprofen/acetaminophen/non-pain medication continued, other previous pain medication discontinued	No pain relief	None reported	3
TN					
Lindström & Lindblom, 1987 ¹⁷⁸	TN (several attacks daily, duration of illness 5–19 years) n = 12 patients on carbamazepine before trial, none experienced complete pain relief	Crossover, double blind, 2 x 2 weeks postoperative tocainide vs. carbamazepine	Tocainide and carbamazepine equally effective	Tocainide: 3/11: pronounced nausea, apical paraesthesias, skin rash	3
Fibromyalgia					
Sörensen <i>et al</i> , 1995 ²⁵⁸	Fibromyalgia (1990 ACR classification), n = 11 (female)	Crossover, double blind, 1 week between treatments: lignocaine i.v. vs. placebo i.v. follow-up 1 week 3/11 on paracetamol	VAS pain intensity: significantly ($p < 0.05$) better after lignocaine vs. placebo	lignocaine: 2/11: nausea + perioral numbness, 1/11: drowsiness, dysarthria, tremor; placebo: 0/11	2

continued

TABLE 31 contd Details of RCTs

Report	Condition; number of patients analysed per group	Design; duration; follow-up	Pain relief	Adverse effects	Quality score
Facial pain					
Marbach & Wallenstein, 1988 ²⁵⁹	Facial pain: n = 28 (women) myofascial pain of the face (n = 15); myofascial pain + osteoarthritis of the TMJ (n = 5); deafferentation neuralgia (n = 4); myofascial pain secondary to deafferentation neuralgia (n = 4); pain duration, 14 months–18 years (mean, 4 years)	Crossover, double blind, 4 x 1 hour at least 1 week between treatments cotton pledgelet lignocaine 4% i.n. vs. cocaine 15% i.n. vs. cocaine 25% i.n. vs. placebo i.n.	No significant effect	Lignocaine: insignificant decrease in blood pressure, 1/21 nausea	3
Cancer pain					
Bruera <i>et al</i> , 1992 ²⁶⁰	Neuropathic pain in cancer patients (numbness/allodynia): direct tumour invasion of nerve plexus. Patients on opioids (mean equivalent daily dose of morphine 231 ± 150 mg) phenytoin, tricyclic antidepressants, corticosteroids already tried, n = 10	Crossover, double blind, 2 x 48 hour lignocaine i.v. vs. placebo i.v.	No difference from baseline, no difference between lignocaine vs. NSAIDs	No significant adverse effects	3
Ellemann <i>et al</i> , 1989 ²⁶¹	Neuropathic pain with allodynia in cancer patients: n = 10; polyneuropathy (n = 7), plexopathy (n = 3) co-analgesics: none (n = 4); paracetamol (n = 2); NSAID (n = 1); opioid (n = 2); opioid + NSAID (n = 1)	Crossover, double blind lignocaine i.v. vs. placebo i.v. 2 x 1 week at least 1 week between treatments, follow-up to see how long effect lasted	No difference between lignocaine and NSAIDs	Lignocaine: 1/10 mild somnolence; placebo: none reported	2
Sjøgren <i>et al</i> , 1989 ²⁶²	Painful bone metastases, duration of pain > 3 m, n = 10; concomitant analgesics: opioid + peripheral (n = 8); epidural morphine (n = 1); peripheral (n = 1)	Crossover, double blind, 2 x 1 week lignocaine i.v. vs. NSAID i.v., 1 week between treatments; if analgesia lasted > 1 week, second infusion only when pain had returned	Mean pain relief no different from placebo, but >10 mm relief: lignocaine 5/10 vs. NSAID 1/10	Lignocaine 4/10: 1, somnolence + nausea; 2, somnolence + circumoral paraesthesias; 3, euphoria; 4, confusion; placebo: none reported	2
Migraine					
Bell <i>et al</i> , 1990 ²⁶³	Acute migraine (common/classic) lignocaine: n = 26; chlorpromazine: n = 24; dihydroergotamine: n = 26 follow-up: lignocaine: n = 17; chlorpromazine: n = 18; dihydroergotamine: n = 19	Parallel group, single blind (patient), 24 hour lignocaine i.v. vs. chlorpromazine i.v. vs. dihydroergotamine i.v.	Post-treatment median score 40 ($p < 0.05?$), change in severity 50% ($p < 0.005?$), complete relief 2/26, incomplete relief 15/26, no relief or worse 9/26, additional medication 11/26	Non-drug-related withdrawals: 19/90; adverse effects: lignocaine: minor 5/17; chlorpromazine minor 4/18; dihydroergotamine: minor 11/19	2

continued

TABLE 31 contd Details of RCTs

Report	Condition; number of patients analysed per group	Design; duration; follow-up	Pain relief	Adverse effects	Quality score
Migraine					
Maizels <i>et al</i> , 1996 ²⁶⁴	Acute migraine with or without aura (IHS), age > 18 years, pain \geq moderate lignocaine: n = 53, placebo: n = 28	Parallel group, double blind, symptomatic and prophylactic medication allowed to continue, follow-up 24 h i.n. vs. placebo i.n. multicentre (2 centres)	50% reduction in VAS pain intensity: lignocaine: 26/53 vs. placebo: 7/28. Rescue medication needed within 4 h: lignocaine: 15/53 vs. placebo: 20/28 ($p < 0.001$)	Lignocaine: minor, local	4
Reutens <i>et al</i> , 1991 ²⁶⁵	Acute migraine (headache lasting > 4 h and < 72 h) n = 25; lignocaine: n = 13, placebo: n = 12	Parallel group, double blind, 2 min infusion, 20 min follow-up, lignocaine i.v. vs. placebo i.v.; paracetamol and/or codeine taken before study	Reduction in VAS pain intensity significant from baseline but not SD from placebo in 20 min	None	2
IHS, International Headache Society					

Four different pain conditions were analysed in 199 patients: peripheral nerve injury, diabetic neuropathy, postherpetic neuralgia and trigeminal neuralgia (Table 31). All three reports of peripheral nerve injury indicated that intravenous lignocaine and oral mexiletine had greater efficacy than placebo, plus a dose response. One study²⁵² reported significant reduction of allodynia after intravenous lignocaine.

Both drugs were also effective in diabetic neuropathy but, in one study,²⁴⁹ the evidence was weak and confined to certain subgroups (patients who had burning, stabbing and heat sensations) in a *post-hoc* analysis. Two studies^{253,254} showed significant relief of dysaesthesia. One study²⁵⁴ reported a significant effect of lignocaine lasting for 8 days.

Lignocaine was also reported to be superior to placebo but inferior to morphine in a study of postherpetic neuralgia.²⁵⁶

Compared with placebo, mexiletine was without effect in dysaesthetic pain following spinal cord injury. Tocainide was comparable to carbamazepine in trigeminal neuralgia.¹⁷⁸

In two reports,^{251,252} dose response was studied, and plasma concentrations of lignocaine were also measured. The minimum effective lignocaine concentration was 1.5 mg/l; this was achieved with doses of 2–5 mg/kg infused over 30–60 minutes.

Dose response was not studied with mexiletine. Dose escalation would indicate that 750 mg daily would provide better analgesia²⁶⁶ and more adverse effects²⁴⁹ compared with 450 mg daily.

Fibromyalgia

Eleven women were studied.²⁵⁸ Pain relief was significant compared with placebo at the end of the infusion and for 15 minutes afterwards. In three of the four responders who had greater than 50% pain relief, the effect lasted for 4–7 days. Tender-point thresholds were not affected. Effects on sleep were not studied.

Facial pain of mixed origin

In one trial in which 28 women with mainly myofascial pain were studied,²⁵⁹ cocaine, 108 mg, administered intranasally resulted in significantly better pain relief compared with placebo, whereas lignocaine was without effect as was a higher dose of cocaine, 180 mg.

Cancer-related pain

Three studies^{260–262} compared intravenously administered lignocaine, 5 mg/kg with saline, in cancer-related pain. Most patients (26/30) were on regular analgesics (strong opioids, NSAIDs, or both). Each study examined a different pain state: pain due to bony metastases,²⁶² chemotherapy-induced polyneuropathy or radiotherapy-induced plexopathy,²⁶¹ and tumour invasion of the nerve plexus.²⁶⁰ Lignocaine had no significant effect.

Migraine

Three reports examined the effect of lignocaine in migraine in 196 patients. Chlorpromazine was significantly better at relieving acute migraine than either intravenous lignocaine (1 mg/kg) or dihydroergotamine.²⁶³ In another study, lignocaine was no better than saline.²⁶⁵ Intranasal lignocaine (20–80 mg) provided significantly better pain relief than saline.²⁶⁴

Adverse effects

In all but one study,²⁶³ ECG was monitored continuously during lignocaine infusions. No arrhythmias were noted. A total of 134 lignocaine infusions were given. Adverse effects were experienced by 21 patients; these were usually minor (light-headedness, somnolence, nausea and perioral numbness). However, the infusion had to be discontinued in five patients and these were considered major adverse effects. A total of 75 patients received an infusion of 5 mg/kg over 30–45 minutes. Of these patients, 16 had minor and three major adverse effects. In two patients, the plasma concentrations were measured at the time of discontinuation. Excessive sedation occurred at 2.42 mg/l and nausea at 2.62 mg/l. One adverse effect was reported during the 100 saline infusions.

Adverse effects were reported in 16 of the 85 patients who were given mexiletine. Adverse effects were dose-related. In three patients they disappeared after the dose was decreased from 10 mg/kg to 8 mg/kg,²⁵³ and in another report,²⁵⁵ adverse effects were reported with the low dose (225 mg/day) by 4/47 patients, with the intermediate dose (450 mg/day) by 1/41 patients, and with the high dose (675 mg/day) by 7/21 patients. The adverse effects were considered mild and they did not necessitate withdrawal from the study. Tocainide caused pronounced nausea in one and apical paraesthesias in another of the 12 patients. A third patient discontinued because of skin rash.

Comment

These results show that sodium channel blockers can reduce pain due to nerve damage. In peripheral nerve injury, diabetic neuropathy or post-herpetic neuralgia, lignocaine was effective at

plasma concentrations of 1.5–5 mg/l. The evidence was strongest in pain due to nerve injury, where the decrease in pain intensity was 40–60%. Oral mexiletine, 750 mg daily, was also effective in these conditions. Allodynia and dysaesthesia were also alleviated. Mexiletine, 450 mg, had no effect in central pain due to spinal cord injury, perhaps because the effects are mainly confined to the peripheral nerves and the dorsal root ganglion. Tocainide has been reported to have caused serious haematological adverse effects including several deaths and should not be used.

Intractable cancer pain is another condition where local anaesthetic type drugs have been advocated.²⁶⁷ The result of this review is quite unequivocal. Intravenous lignocaine, 5 mg/kg, was without effect. It does not give more relief when combined with high doses of opioids and NSAIDs.

Intranasal lignocaine in migraine was significantly more effective than placebo.²⁶⁴ The NNT for the reduction of pain to mild or none was 3 for intranasal lignocaine compared with placebo. The NNT for the same endpoint with subcutaneous sumatriptan was 2.²⁶⁸ Headache recurred within 24 hours in 42% of patients after intranasal lignocaine and between 30% and 48% after subcutaneous sumatriptan.

The long-term analgesic effects of intravenous lignocaine were not systematically studied. Only two studies^{254,258} reported that pain relief could last for several days. It is not known if subsequent infusions provide longer relief. Long-term systemic administration of lignocaine is not practical, and no controlled studies have been undertaken on either its efficacy or adverse effects in long-term use.

It is not known if there are patients who benefit from lignocaine who do not benefit from mexiletine. There seems to be little point in using lignocaine infusion to predict response to mexiletine if response can be gauged by taking mexiletine alone.

The encouraging signal from this review is that a difficult subgroup of neuropathic pains can be helped. Delivering and optimising that benefit will require new approaches. In the meantime, the important message is that all neuropathic pains do not necessarily respond identically.

Chapter 17

Topical NSAIDs in chronic painful conditions

Despite their licensed status, there is scepticism that topical NSAIDs have any action other than as rubefacients.^{269,270} This systematic review was undertaken to examine the evidence that topical NSAIDs are safe and effective and, if possible, to determine which ones were the most effective in chronic painful conditions.

Methods

Reports were sought of RCTs of topical NSAIDs in which pain was an outcome. Reports were included which compared topical NSAIDs with placebo, with another topical NSAID, or with an oral NSAID. A number of different search strategies in both Medline (1966–May 1996) and the Oxford Pain Relief database (1950–94) were used to locate reports, using individual drug names, together with the terms, ‘administration, topical’, ‘gels’, ‘ointments’, ‘aerosols’, ‘cream’, and combinations of these, without restriction to English language. Additional reports were identified from the reference lists of retrieved reports and from review articles. Medical librarians and medical directors of 12 pharmaceutical companies which make NSAIDs were asked for reports of RCTs of these products, including any unpublished reports.

RCTs of NSAIDs in chronic arthritic, rheumatic or associated conditions with pain as an outcome were included in the analysis, but not acute traumatic conditions, vaginitis, oral or buccal conditions, thrombophlebitis or experimental pain.

Two of the reviewers screened reports to eliminate those which had no pain outcomes, which were definitely not randomised, or were abstracts or reviews. The methodological quality of each trial was assessed by all reviewers using a validated scale.⁷

Information about treatments and controls, 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 withdrawals due to adverse events was extracted from each report by all six authors meeting to concur.

A clinically relevant outcome was defined as at least 50% pain relief. Information was extracted in dichotomous form for analysis. The denominator was taken as the number of patients randomised (an intention-to-treat analysis). A hierarchy of measures was used which approximated to the following, in this order:

- (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.

For chronic conditions, we took the effectiveness measure nearest to 14 days after start of treatment.

Results

Placebo-controlled trials

There were 11 studies including 1067 patients, of whom 532 received active topical NSAID and 535 a non-NSAID placebo. Two studies used diclofenac plasters, two salicylates, two a flufenamate/salicylate combination and one each diclofenac, felbinac, indomethacin, ibuprofen and flurbiprofen (see *Figure 25*). Two reports had quality scores of 2, five of 3 and four of 4.

Combining the results across all these studies (since there was no *a priori* indication that one was better than another) gave a relative risk of 2.1 (1.5–2.8) and an NNT of 3.1 (2.7–3.8) for more than 50% relief compared with placebo at 2 weeks.

Adverse events were rare and there was little difference between topical NSAIDs and placebo. With active topical NSAID in 532 patients, 19 had a local reaction, five a minor systemic reaction and one an adverse event that necessitated drug withdrawal. For the 535 patients given placebo, the respective figures were none, three and none.

Active-controlled studies

Seven studies with over 500 patients studied two different topical NSAID regimens, two studied

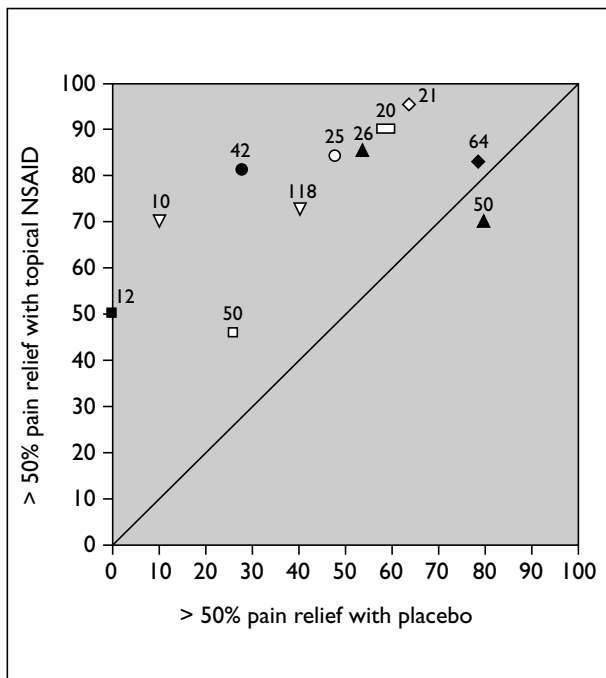


FIGURE 25 Placebo-controlled studies of topical NSAIDs in chronic pain conditions (□ trolamine salicylate, ◇ felbinac, ○ diethylamine salicylate, ▲ diclofenac plaster, ■ indomethacin, ● ibuprofen, □ diclofenac, ▽ flufenamate + salicylate, ◆ flurbiprofen)

topical NSAID against oral aspirin or ibuprofen, and one added a topical NSAID to half the oral dose of ibuprofen. In these ten trials, quality scores were 1 in one trial, 4 in two, 3 in three and 4 in two.

There was no difference between topical piroxicam gel and ibuprofen, 1200 mg daily, or topical trolamine salicylate against aspirin, 2600 mg daily.

The single arms of these trials have been taken and plotted together with those from the placebo-controlled trials (*Figure 26*). No extractable information was available from two studies.

Discussion

Most RCTs of topical NSAIDs have been for the treatment of acute traumatic pain conditions such as soft tissue injuries, sprains and strains; over 50 such studies have been reported. Treatment of chronic arthritic or rheumatic conditions or conditions such as tendinitis with topical NSAIDs is rare, and there are few cases where the same NSAID has been tested in more than one study.

Despite this, the analgesic effect of the topical preparations was good. Overall, the preparations could clearly be distinguished from placebo in

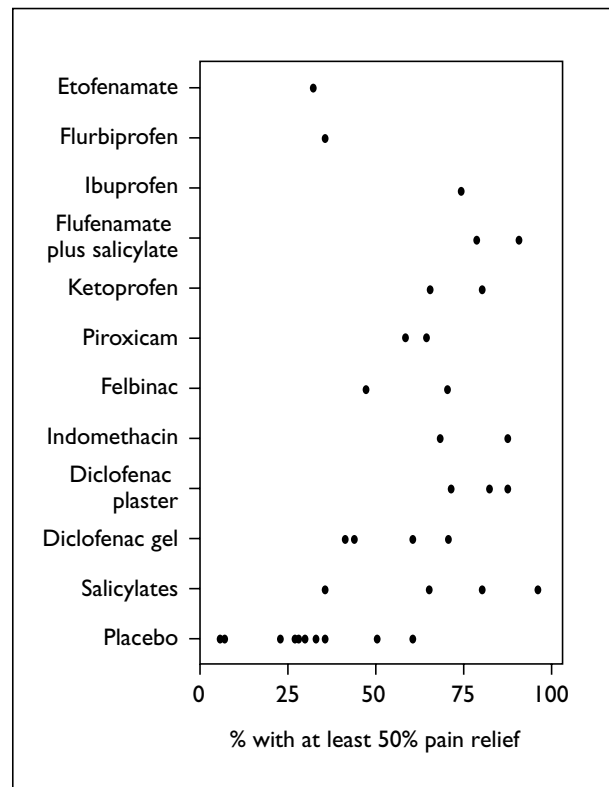


FIGURE 26 Plot of proportion with at least 50% pain relief with the different drugs in chronic conditions

chronic pain conditions, and they produced a combined NNT of 3.1, similar to results obtained with oral, intramuscular and sublingual NSAIDs given for postoperative pain relief, and to anti-convulsant drugs in chronic pain, using a similar criterion of greater than 50% pain relieved. An NNT of 3 indicates that for every three patients treated with an NSAID, one would have at least half the pain relieved who would not have had they been given placebo. Analysis of oral NSAID or simple analgesics has not been undertaken but experience would suggest they are unlikely to be more effective. Indeed, two studies, one of which had over 100 patients per group, failed to detect any difference between topical NSAID and oral aspirin or ibuprofen. In a third study, topical NSAID was substituted for half the daily ibuprofen dose.

Results of the meta-analysis show that, in chronic pain, local and systemic adverse effects were rare and their incidence could not be distinguished from placebo preparations.

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. Plasma concentrations of ibuprofen in eight

subjects, after 400 mg was taken orally, peaked at 39 µg/ml, while the peak after 300 mg was applied to the skin was 0.6 µg/ml.²⁷¹ Concentrations of about 0.3 µg/ml were found²⁷² after application of 400 mg topical ibuprofen cream. Topical application of ibuprofen resulted in significant tissue concentrations in deep tissue compartments, more than enough to inhibit inflammatory enzymes.^{272,273}

The question has been raised of whether using oral NSAIDs is worse than the disease for some patients.²⁶⁹ Information from an on-going systematic review of evidence from studies with a variety of structures indicates the prevalence of ulcer at 2% after oral NSAID, an incidence of upper gastrointestinal bleeding of about 0.2%, with one patient in ten with a gastrointestinal bleed or perforation dying. For the average patient with rheumatoid arthritis taking oral NSAIDs, the attributable risk of going to hospital with gastrointestinal problems is 1.3–1.6% annually, and the risk of death is 0.15%.²⁷⁴

Ten million people aged 60 years and older live in England and Wales. Osteoarthritis is the most common form of arthritis; the number of sufferers is likely to increase from about 1.1 million now to 1.3 million by 2020, as the population ages and the number of people aged over 60 years, who are most likely to suffer from the disease, increases. Rheumatoid arthritis affects 2% of men and 5% of women over the age of 64 years. In the UK, some 20–24 million NSAID prescriptions, 5% of the total of all NHS prescriptions, are written each year for osteoarthritis at a cost of over £180 million.

A detailed cost-minimisation analysis comparing topical versus oral NSAIDs for the treatment of mild osteoarthritis of the superficial joints²⁷⁵ indicates that topical NSAIDs can reduce overall costs of treatment significantly. The study assumed that topical NSAIDs were not associated significantly with upper gastrointestinal bleeding; this is supported by a recent report.²⁷⁶ Using a simple decision tree to represent the two treatment choices and the consequences, clinical and economic, that ensue from each decision, the authors

generated an expected cost for 3 months' treatment of £89 for oral ibuprofen and £54 for topical piroxicam gel, based on equal efficacy of these two treatments.²⁷⁵ The calculations did not include co-prescribing of misoprostol or acid-suppressing medicines with the oral NSAIDs, which would increase savings even more.²⁶⁹

More general use of topical NSAIDs in both acute and chronic pain conditions could give good relief with minimal adverse effects to many people – better quality care at lower cost.

Comment

There is a *prima facie* case that topical NSAIDs are effective in chronic painful conditions, particularly single arthritic joints. The problem is that most of these trials were performed with preparations unavailable in the UK, and without either sufficiently large numbers, or numbers of trials, to make the case with certainty. Indeed, most only addressed the first 2 weeks or so of treatment, so that information on adverse events, although not different from placebo, is of little value considering that the topicals would need to be used for many weeks or many months.

However, there is also a suggestion that, despite their higher prescription cost than oral NSAIDs or simple analgesics like paracetamol, effective topical NSAIDs may have a cost-benefit to the health service through lower rates of serious complications like gastrointestinal bleeding. Benefit to patients is obvious through reduction in harm, and studies show that gastrointestinal complications with topical NSAIDs are no higher than in the population at large.

What is needed is at least one, large, RCT using the topical NSAID likely to produce the greatest benefit (probably ibuprofen, though other candidates are possible). This study would probably have two phases – an early more intense phase to establish effectiveness, and a longer, on-going phase to determine both efficacy and safety.

Chapter 18

Topical capsaicin

Capsaicin is an alkaloid from chillies that first entered European knowledge after Columbus' second voyage to the new world in 1494. It has been a feature of pharmacopoeias for many years.

Recent interest concerns the use of topical capsaicin as an analgesic for a variety of conditions where pain may not be responsive to classical analgesics. There is evidence that capsaicin can deplete substance P in local nerve sensory terminals. Substance P has been thought to be associated with initiation and transmission of painful stimuli, as well as a number of diseases – arthritis, psoriasis and inflammatory bowel disease. This has given topical application of capsaicin some degree of logic – remove the neurotransmitter, and remove the pain.

Does it work?

Zhang and Li Wan Po⁷⁵ searched the literature for capsaicin papers using a sensitive strategy. They sought reports of clinical investigations. Only information from randomised, double-blind and placebo-controlled studies were used for quantitative analysis by clinical condition.

Results for the 13 trials that fulfilled these criteria and where there were extractable data are shown as a L'Abbé plot (Figure 27). Each symbol represents the proportion of patients in each trial reaching some clinical end-point for benefit and the number next to it the number of patients given topical capsaicin. Capsaicin results are plotted against placebo results. Points lying between the line of equality and the capsaicin axis are trials showing benefit. This plot is a simple representation of how similar or dissimilar trial results were found to be.

Diabetic neuropathy

Four trials reported on the use of capsaicin cream, 0.075%, applied four times daily for 4–8 weeks in diabetic neuropathy in a total of 144 patients treated with capsaicin and 165 with placebo cream. The end-point was a physician global assessment of pain relief. Clinical improvement was pain completely gone, much better or better (and not; no change, worse, or much worse).

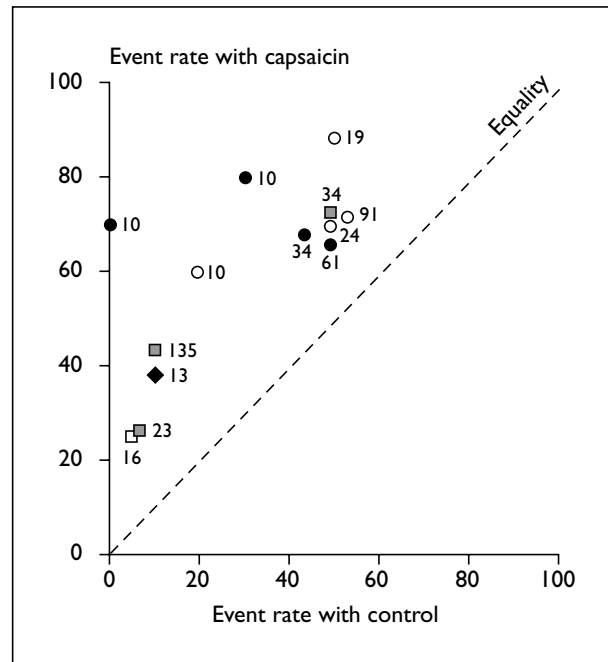


FIGURE 27 L'Abbé plot for topical capsaicin in various conditions, with numbers of patients shown for each point (○, diabetic neuropathy; ●, psoriasis; ◆, post-mastectomy pain; ■, osteoarthritis; □, PHN)

In all, 105/144 (73%) patients responded with capsaicin compared with 81/165 (49%) patients given placebo. The relative risk favouring capsaicin was 1.5 (95% CI, 1.2–1.8) and the NNT was 4.2 (2.9–7.5).

For every four patients treated with topical capsaicin, one would have had pain of diabetic neuropathy relieved who would not have had they been treated with placebo.

For comparison, oral anticonvulsant therapy for diabetic neuropathy in 66 patients treated in two trials yielded an NNT of 2.5 (1.8–4.0).⁷⁶

Osteoarthritis

Three trials reported on the use of capsaicin cream (0.025% in two, and 0.075% in one) four times daily for four weeks in osteoarthritis. The end-point was articular tenderness or the physician's global assessment of pain relief.

A total of 87/192 (45%) patients responded with capsaicin compared with 30/190 (16%)

with placebo. The relative risk favouring capsaicin was 2.9 (2.0–4.1) and the NNT was 3.3 (2.6–4.8).

For every three patients treated with topical capsaicin, one would have had the pain of osteoarthritis relieved who would not have had they been treated with placebo.

Postherpetic neuralgia

Only a single trial that fulfilled the inclusion criteria was available. Use of a 0.075% cream three or four times daily for 6 weeks resulted in pain relief in 4/16 patients with capsaicin compared with 1/16 patients with placebo. Relative risk was 14.0 (0.5–32) – no significant improvement.

Postmastectomy pain

A single trial of 0.075% cream four times daily for 6 weeks resulted in pain relief in 5/13 patients with capsaicin compared with 1/10 patients with placebo. Relative risk was 3.9 (0.5–28) – no significant improvement.

Psoriasis

Four trials reported on topical capsaicin, 0.025%, four times daily for 6–8 weeks in psoriasis. Psoriasis was rated as to the degree of itching, scaling and erythema. The end-point was much better or better rating of overall appearance.

In all, 78/115 (68%) patients responded with capsaicin compared with 55/130 (42%) with

placebo. The relative risk favouring capsaicin was 1.6 (1.3–2.0) and the NNT was 3.9 (2.7–7.4).

For every four patients treated with topical capsaicin, one had symptoms of psoriasis relieved who would not have if they had been treated with placebo.

Comment

The authors of the review suggested that blinding of trials of capsaicin might be difficult because of its irritant effects when applied to the skin. There also appear to be suggestions from some of the reports that skin irritation wears off with time, while analgesic effects may improve with time.

How much the placebo effect may influence the results is uncertain. These are difficult clinical conditions, and patients used the creams for up to 8 weeks. Variability in the results in placebo groups can be seen from *Figure 27* to be from 0% to 50% of patients on placebo getting benefit. All the trials showed benefits over placebo, although not all trials were themselves statistically significant.

The numbers of patients treated in these studies was not great but that is not unusual in these difficult clinical conditions.⁷⁶ The review did not include results of adverse effects, which is a shame, since any treatment choice balances the probability of benefit and the risk of harm.

Chapter 19

Psychological approaches

Introduction

Effectiveness and cost-effectiveness

Psychologically-based treatments for chronic pain are becoming more common. This has raised a number of important questions regarding the effectiveness of various treatments and combinations and deliveries of treatment. As with all pain treatments, regulation of standards of treatment and competent studies of the cost-effectiveness of these treatments are rare. Summarised here is the current evidence on what is effective, what is dubious or of no apparent efficacy, and what requires further investigation.

Study quality

Addressing these questions requires attention to the following features of studies: source of information (country and setting of study, date of publication, researcher affiliations); trial design (randomisation procedures; treatment modes; length of treatment and of follow-up; treatment validity; therapist qualities; credibility of and satisfaction with treatment; adherence to treatment methods); and participants (details of sampling; inclusion and exclusion criteria; personal, socio-economic and medical characteristics). In addition, the levels of treatment and measurement of outcome require detailed examination for comparison across studies to have any validity.

Summary review

Since a full systematic review and meta-analysis will not be complete until September 1997 (coding of all the variables above for each treatment component and for each measure has been piloted and is, at the end of 1996, in process using three independent raters), the complete set of studies is described here, with ratings of quality (using the criteria developed by Jadad and colleagues⁷), tabulated according to study population and initial meta-analysis. Two model studies are described as a guide to appropriate treatments (not a comprehensive set) and achievable outcomes; both involved patients (rather than volunteers) and were carried out within orthodox treatment facilities.

Psychologically-based pain treatments

Aims of treatment

Unlike most other treatments for chronic pain, psychologically-based treatment aims to restore function and psychological hardiness despite continuing pain, since the covariance of pain, disability and mood is low.^{277,278} These aims should imply extensive and well-developed measurement of that function, but the field has been dominated by self-report measures, some of dubious utility. Measurement is described and discussed later (see page 101). Similar gains in psychological health and function may result from non-psychological treatments even in the absence of total pain relief: these, however, are rarely measured in the context of the assumption that abolition or substantial reduction in pain is a necessary condition for restoration of function. It is probable that widespread consideration of such psychological contributions to recovery as placebo, and thereby discountable, effects⁴¹ contributes to the reluctance to assess psychological variables in studies of medical or psychological intervention.

Psychological intervention

Psychological intervention is usually one component of a multicomponent rehabilitative programme; however, the other components, such as exercise, graded activity increases, drug reduction, and abandonment of unnecessary aids, are underpinned by behavioural principles,^{277,279} and to a lesser extent by understanding of cognitive psychology. The interventions described here aim to improve patients' activity level and to reduce maladaptive pain behaviours and drug intake, mainly by operant methods; to improve control over pain and its adverse effects, mainly by relaxation techniques; to enhance maintenance of treatment gains, by operant and cognitive methods; and to mitigate negative mood and revise unhelpful beliefs by cognitive methods.

Operant methods

The basis of the use of operant methods in pain was originally stated by Fordyce,²⁷⁷ and has changed little since. It is anchored in the gate-control model understanding of the modulation

of pain transmission at a spinal level, not least by cortical influences.

“The thesis ... is not that pain is originally produced by operant conditioning ... but that much of the behavior occurring subsequent to presentation of a presumed noxious stimulus may be accounted for and modified by principles of learning, whatever the original cause of pain.”²⁷⁷

Thus operant methods are used to increase desired behaviours (such as a range of activities, correct posture and movement, talking about subjects other than pain and suffering) and to decrease or extinguish unhelpful habits (such as excessive rest, overuse of drugs in relation to their limited benefits, frequent verbal and nonverbal expression of pain). Maximum effectiveness requires careful observation of antecedents and consequences of target behaviours, ingenuity in manipulation of these variables, and high levels of team consistency in the delivery of socially-rewarding responses to patients, most usually attention, interest and praise.

Behavioural methods

While incorporating operant methods, the umbrella of behavioural techniques includes three of particular relevance here, all with extensive evidence of effectiveness in mainstream psychological treatment. These are:

- (i) the use of graded exposure methods to increase engagement in feared activities and to reduce fears (for application to patients with pain, see Gil and colleagues²⁸⁰)
- (ii) the use of relaxation in modifying somatic responses to feared situations or events and enhancing a sense of control
- (iii) the practice of clear communication and assertion skills.

Cognitive methods

Cognitive methods originally relied on those of demonstrated utility in acute pain, in clinical or experimental situations, in which manipulation of attention (e.g. distraction) attenuated pain.²⁸¹ However, there are considerable problems with the experimental models, particularly in the light of information processing theories,²⁸² and since evidence of their efficacy in chronic pain, other than headache, is wanting.

The related dimensions of control and coping have been investigated; however, control emerged as a complex and multidimensional variable, and the concept of coping suffered from weak theoretical roots and overextension to describe processes of

appraisal, strategies enacted, and their consequences. Self-efficacy²⁸³ offered a version of control of clinical relevance; of greatest utility is the concept and measure of catastrophising,^{284,285} the belief system or thinking habit in which the worst is expected and negative predictions made about its impact on the person.

Recently, the cognitive model of depression of Beck and his school²⁸⁶ has been imported, offering theory and therapies superior to those arising from imported psychiatric models.^{287,288} Turk and colleagues²⁸⁹ describe

“...the underlying assumption that affect and behavior are largely determined by the way in which the individual construes the world. Therapy is designed to help the patient identify, reality-test, and correct maladaptive, distorted conceptualisations and dysfunctional beliefs.”²⁸⁹

Cognitive-behavioural therapy

The methods above are recombined variously into what is usually described as cognitive-behavioural therapy (CBT). While it is not necessarily true that compound treatments bring all the benefits of their constituent elements,²⁹⁰ cognitive change may reliably result from behavioural change (for instance, the change in beliefs that increased activity will cause harm when this prediction is repeatedly disconfirmed), as may mood improvement (reduction in depression from engagement in pleasant activities²⁹¹). Conversely, behavioural change may result from altered beliefs (such as learning that increased pain on movement may signal reversible muscle tension rather than irreversible tissue damage). Combining behavioural and cognitive methods attempts to capitalise on the beneficial changes associated with each.

Reviews and analyses

Reviews

Published reviews

A total of 12 selective reviews of treatment gains from largely uncontrolled studies^{292–303} and two of maintenance^{290,304} produced conclusions of guarded optimism about the efficacy of behavioural methods for attenuating pain and restoring function in chronic pain patients, with criticism of measurement techniques, brief or no follow-up period, and extravagant programme content.

Meta-analyses

Two meta-analyses^{305,306} have been published and at least one more is in progress,³⁰⁷ although none of

these is restricted to RCTs. Malone and Strube,³⁰⁵ who included headache and dental pain in their 109 studies, found outcome effect sizes greater than 1.5 for mood; between 1.5 and 1.0 for activity level, drug use and subjective symptoms including some pain report; and less than 1.0 but greater than 0.5 for pain intensity and self-rated improvement. Analysed by treatment, largest effect sizes (over 2.0) were obtained for autogenic and hypnotic treatments (largely relevant to headache and dental pain populations); between 1.5 and 1.0 for multimodal treatments, and less than 1 but greater than 0.5 for biofeedback, cognitive treatment, relaxation, and operant treatment, in descending order.

A more rigorous meta-analysis by Flor and colleagues³⁰⁶ entered 65 studies published between 1960 and 1990 which met criteria concerning the quality of psychological treatment within a multidisciplinary pain clinic. Headache was excluded. The aim was to estimate the size of change within groups across treatment, or between treatment and control group, according to modality of outcome measure and characteristics of the study population and of the study itself. Most treatments were multimodal; half the studies were on inpatient treatment, 28% on outpatient treatment, and 13% were mixed. Mean age of patients was 45 years, and half were male; mean pain duration was 85 months, the median being 60 months. Over half had undergone at least one surgical operation for pain and 34% were working. Because of the likelihood of correlation between multiple measures within the same study, the authors calculated a mean effect size for each type of dependent measure and for each study, separately for up to 6 months follow-up and for more than 6 months. Overall effect sizes for studies as a whole were about 1.5 (SD, 1.5) for up to 6 months and 1.3 (SD, 1.1) for over 6 months. These are equivalent to 60% and 55% changes, respectively. Effect sizes for measures of control groups were rarely significantly different from zero, although differences between the scores of treated patients and controls narrowed with longer follow-up; studies with random assignment of patients to treatment conditions tended to have smaller between-group effect sizes.

Among the various dependent measures, greatest improvements appeared in self-report of pain (pre- to post-treatment effect size of 1.6), followed by behavioural measures (not further specified: effect size of 1.2); among the specific measures, drug reduction and return to work both showed a change of over 60%, compared to under 30% in controls; reduction in health care use was more

modest. None of the population characteristics tested (age, duration of pain, litigation and compensation) emerged as a predictor of study effect size; neither did length of treatment or attrition rate (recruitment to follow-up), although sample sizes for all comparisons but length of treatment were relatively small.

Ratings of quality of study by the reactivity of measures, by their reliability and validity and by the possibility of bias in assessment, revealed widespread use of reactive and poor quality measures: there was a slight association of effect size with reactivity of measures. The authors concluded that

“...Overall the results of this meta-analysis provide support for the conclusion that multidisciplinary pain clinics are efficacious. Even at long-term follow-up, patients who are treated in such a setting are functioning better than 75% of a sample that is either untreated or that has been treated by conventional, unimodal treatment approaches.”³⁰⁶

It was partly the widespread dissatisfaction expressed in reviews of the apparent lack of specificity of effects, and partly the need to economise on resources expended in treatment, which encouraged the execution of component dismantling studies and of other RCTs, reviewed below.

Search strategy for the current review

A total of 23 studies were known to the authors of this review; two were supplied pre-publication by the authors who were aware of this review; two appeared in *Current Contents* after the search had been completed. A liberal search strategy was used, casting a wide net over the available resources. Four related electronic databases were searched (Medline, Embase, Psychlit and Social Science Indices). Following Jadad and colleagues,⁷ core search terms were included, with psychological treatment terms liberally added to these. The search was deliberately expansive. Only 60% of those RCTs entered into the final analyses were found in Medline. Only one study was recorded on the Social Sciences Indices which was not recorded elsewhere. Work is in progress to identify the most efficient search strategy, working across related databases. Eight studies were added later: two appeared after the initial searching was completed, and a further six were gained from reference sections of target papers. Many of these are studies of rheumatoid arthritis, pelvic pain and chest pain populations. Papers were included if they met the following criteria:

- (i) publication in a refereed journal

- (ii) a statement was included that patients were randomly allocated to treatment (even if methods were questionable)
- (iii) any chronic pain other than headache was included
- (iv) treatment included a stated primary psychological component.

In all, 35 studies met these criteria and were included in the quality review and meta-analytic process. A complete list of these studies is given after the main reference list at the end of this document.

Studies included

Component dismantling studies

Within treatments which include deliberate psychological intervention, as with those without,

the means by which psychological gains are made remain obscure. It is clear from component studies that it is neither the case that specific interventions result in gains in the relevant variables only, and conversely that those gains are only made by the specific intervention, nor that gains are altogether non-specific with respect to intervention;³⁰⁸ nor is it a simple matter of enhancing patient confidence or control. However, interpretation of studies of psychological treatment is problematic.

As is clear from *Tables 32* and *33*, the majority of studies use component-dismantling approaches to address the question of the value of a particular component when added to others, or to a basic treatment. There is a fundamental problem

TABLE 32 RCTs: low back pain or mixed chronic pain or specific site: treatment content

Authors	Date	Study
Flor & Birbaumer	1993	Biofeedback/relaxation vs. CBT vs. routine medical/physical treatment
Kerns <i>et al</i>	1986	Operant vs. CBT vs. waiting list control
Linton & Gotestam	1984	Applied relaxation vs. applied relaxation + operant vs. waiting list control
Linton <i>et al</i>	1985	Operant + routine treatment vs. relaxation + operant + routine treatment vs. waiting list control
Moore & Chaney	1985	CBT + spouse vs. without spouse vs. waiting list control
Newton-John <i>et al</i>	1995	CBT vs. biofeedback/relaxation vs. waiting list control
Nicholas <i>et al</i>	1991	Operant vs. cognitive, both + vs. – relaxation vs. attention
Nicholas <i>et al</i>	1992	CBT vs. attention control vs. waiting list control
Peters & Large	1990	In-patient vs. out-patient CBT vs. waiting list control
Peters <i>et al</i>	1992	Follow-up in-patient vs. out-patient vs. waiting list control to 1 year
Turner	1982	Relaxation vs. CBT vs. waiting list control
Turner & Clancy	1988	Operant vs. CBT vs. waiting list control
Turner <i>et al</i>	1990	Behavioural with vs. without exercise vs. exercise
Turner & Jensen	1993	Cognitive vs. relaxation vs. cognitive + relaxation
Vlaeyen <i>et al</i>	1995	Operant vs. operant + cognitive vs. operant + relaxation vs. waiting list control
Williams <i>et al</i>	1996	In-patient vs. out-patient CBT vs. waiting list control to 1 year
Altmaier <i>et al</i>	1992	Standard rehabilitation + vs. – psychological component
Pilowsky <i>et al</i>	1995	CBT + amitriptyline vs. support + amitriptyline
Puder <i>et al</i>	1988	CBT immediate vs. CBT delayed
Vlaeyen <i>et al</i>	1996	Fibromyalgia: cognitive + education vs. education + discussion
Spence <i>et al</i>	1989	Upper limb: individual CBT vs. group CBT
Spence <i>et al</i>	1991	Upper limb: follow-up to 2 years
Spence <i>et al</i>	1995	Upper limb: biofeedback/relaxation vs. applied relaxation
Peters <i>et al</i>	1991	Pelvis: treatment for psychological problems vs. routine treatment
Cott <i>et al</i>	1992	Chest: individual CBT vs. group CBT vs. attention vs. waiting list control

TABLE 33 RCTs: rheumatoid arthritis or osteoarthritis: treatment content

Authors	Date	Study
Keefe <i>et al</i>	1990a	Osteoarthritis knee: pain coping skills vs. education vs. routine treatment
Keefe <i>et al</i>	1990b	Osteoarthritis knee: follow-up to 6 months
Keefe <i>et al</i>	1996	Osteoarthritis knee: pain coping skills with vs. without spouse vs. education
Kraimaat <i>et al</i>	1995	Rheumatoid arthritis: CBT vs. occupational therapy vs. waiting list control
Parker <i>et al</i>	1988	Rheumatoid arthritis: CBT vs. attention control
O'Leary <i>et al</i>	1988	Rheumatoid arthritis: CBT vs. self-management
Bradley <i>et al</i>	1987	Rheumatoid arthritis: CBT vs. social support vs. waiting list control
Appelbaum <i>et al</i>	1988	Rheumatoid arthritis: CBT vs. symptom monitoring
Radojevic <i>et al</i>	1982	Rheumatoid arthritis: operant with vs. without family vs. education with family
Strauss <i>et al</i>	1986	Rheumatoid arthritis: group psychotherapy vs. group assertion + relaxation

with the implicit model of treatment additivity and thus with the additivity of treatment effects. Linton²⁹⁰ was critical of the apparent assumption that 'more equals better' in programme content (the 'everything-but-the-kitchen-sink' approach), and emphasised the problems of cost, the potential for adverse effects on adherence, and the possibility that the benefits of effective components might be undermined by ineffective components. This is a particular issue when much is made of the integration of treatment components and the efficient functioning of an interdisciplinary (interactive) rather than merely multidisciplinary (additive) team.³⁰⁹ Nevertheless, the question of whether elements of treatment are redundant or even counter-therapeutic are of clear practical importance, with particular relevance to questions of cost-effectiveness.

Addressing the studies for their practical, rather than theoretical, contributions to knowledge, the major problems for comparison of treatments are:

- (i) the lack of treatment details and differences in treatment levels
- (ii) the lack of shared measures, with underuse of appropriate, objective and/or standardised assessment procedures and instruments.

Coding categories

Coding categories were developed by discussion between the authors and by reference to the literature on effectiveness of psychological therapies.^{310,311} Particular attention was paid to issues of process, largely neglected in the review literature in this field. This generated five sets of coding (see Box on page 113), concerning source of information, trial design, participants, treatment components and effect size

(for each measure). A pilot set was attempted on a sample of studies and the results discussed. The revised codings were completed by the authors of this chapter of the review independently and a consensus achieved by discussion.

Source of information

The source of information will not be addressed in detail here: almost all studies were carried out in university-affiliated pain clinics, the majority in the USA or Canada, with the exceptions of Vlaeyen *et al* (1995; 1996), Kraimaat *et al* (1995), and Peters *et al* (1991) in The Netherlands; Linton and Gotestam (1984) and Linton *et al* (1985) in Sweden; Flor and Birbaumer (1993) in Germany; Nicholas *et al* (1991; 1992), Spence (1989; 1991), Spence *et al* (1995), Newton-John *et al* (1995), and Pilowsky *et al* (1995) in Australia; Peters and Large (1990) and Peters *et al* (1992) in New Zealand; Williams *et al* (1996) in the UK. A full list of these studies is given after the main reference list at the end of this document (see page 134).

Participants

No details of participants are discussed here other than to note that the populations in the studies by Turner and Clancy (1988) and O'Leary and colleagues (1988) were, in part, volunteers recruited through advertisements and who were not at the time seeking any therapeutic intervention. In general, as indicated by Turk,³¹² patients of multi-disciplinary pain centres tend to have the most recalcitrant pain problems: most have undergone surgical and other interventions which have not relieved the pain or returned them to adequate functioning; many continue to seek further interventions and to take drugs whose meagre benefits hardly outweigh their adverse effects. There is also

TABLE 34 Quality ratings: low back pain or mixed chronic pain or specific site studies

Study	Randomisation	Credibility	Therapist equivalence	Attrition	Total	Number of patients	Comments
Flor & Birbaumer 1993	1	1	1	1	4	78	
Kerns <i>et al</i> 1986	1	1	1	1	4	28	
Linton & Gotestam 1984	1	1	1	1	4	15	
Linton <i>et al</i> 1985	1	0	?	1	2	28	
Moore & Chaney 1985	1	0	0	1	2	43	
Newton-John <i>et al</i> 1995	1	0	?	1	2	44	
Nicholas <i>et al</i> 1991	1	1	1	1	4	58	
Nicholas <i>et al</i> 1992	1	1	1	1	4	20	
Peters & Large 1990	2	0	0	1	3	85	
Turner 1982	0	0	1	0	1	36	random?
Turner & Clancy 1988	1	1	1	1	4	81	
Turner <i>et al</i> 1990	1	0	1	1	3	96	
Turner & Jensen 1993	1	0	1	1	3	90	
Vlaeyen <i>et al</i> 1995	1	0	?	1	2	71	
Williams <i>et al</i> 1996	2	1	1	1	5	114	
Altmaier <i>et al</i> 1992	1	0	?	0	1	45	
Pilowsky <i>et al</i> 1995	1	0	1	0	2	64	random? 37% attrition
Puder <i>et al</i> 1988	1	0	1	0	2	69	
Vlaeyen <i>et al</i> 1996	1	1	1	1	4	125	
Spence <i>et al</i> 1989	1	1	1	1	4	45	
Spence <i>et al</i> 1995	1	0	1	1	3	48	
Peters <i>et al</i> 1991	2	0	0	1	3	106	treatment unclear
Cott <i>et al</i> 1992	1	0	0	0	1	104	33% drop-out

a reasonably high prevalence of depression of clinical concern, long-term loss of work and of family roles and valued activities, and no evidence of spontaneous recovery from any of these difficulties over the years of chronic pain.

Patients in the 35 studies were described according to the following variables (the number of studies using each is given in parentheses): in the personal domain, age (31), sex (27), marital status/home situation (15), educational level (13), employment status (20), time since last employed (2), job changes due to pain (1), salary/family income (2), receipt of disability compensation/benefit (16), seeking change in disability compensation status (1), pending litigation (6); and in the medical domain, aetiology of pain (7), duration of pain (25), age of onset (2), significance of radiological findings (2), rheumatoid arthritis criteria (10), percentage with

previous pain-related surgery (11), percentage with previous physiotherapy (2), percentage with previous nerve blocks (1), recent hospital inpatient days (1), hospital outpatient visits (1).

Trial design

In *Tables 34* and *35*, studies have been scored using the method of Jadad and colleagues,⁷ adapted for double-blindness (self-evidently not possible in psychological treatments) by substituting the requirements of evidence for equal credibility or expectation of benefit across treatments (patient 'blindness') and treatment type unconfounded with therapist, most usually by the same therapist or therapists administering all treatment conditions (therapist 'blindness'). In some senses, these conditions are more stringent than those of Jadad, since blindness of therapist and patient is assumed, given certain

TABLE 35 Quality ratings: rheumatoid arthritis or osteoarthritis trials

Study	Randomisation	Credibility	Therapist equivalence	Attrition	Total	Number	Comments
Keefe <i>et al</i> 1990a,b	2	1	1	1	5	99	
Keefe <i>et al</i> 1996	1	1	1	1	4	88	
Kraimaat <i>et al</i> 1995	1	0	0	1	2	77	
Parker <i>et al</i> 1988	2	1	1	0	4	83	
O'Leary <i>et al</i> 1988	1	0	?	1	2	33	
Bradley <i>et al</i> 1987	1	1	?	1	3	18	
Appelbaum <i>et al</i> 1988	1	0	1	1	3	53	
Radojevic <i>et al</i> 1982	1	1	1	0	3	59	
Strauss <i>et al</i> 1986	1	0	0	0	1	57	treatment odd

conditions (such as indistinguishableness of active and placebo pill), but rarely checked.

However, there are several other features of psychological treatment which could be considered as important for rating study quality: they concern treatment integrity (issues of therapist skills and competence, related to training, experience and supervision); treatment fidelity (blind rating of audiotapes of treatment sessions to determine that conditions are distinguishable by content); treatment replicability (ideally, treatments are thoroughly manualised); patient adherence during and after treatment to treatment methods; and patient post-treatment ratings of helpfulness of treatments or components of treatment (which can improve estimates of patient 'blindness' to condition). These are not considered further here; all are found in at least one, and usually in several, of the studies in this review.

Treatment content

Treatment content is coded in *Tables 36–38** without any indication of the level of intervention. Within each treatment component (education, relaxation, exercise, behavioural management, cognitive therapy, problem solving, social/family involvement, generalisation and maintenance) there are issues of the number of hours of therapeutic contact; the requirement for homework and issues of adherence during and between treatment sessions; the level of detail and appropriateness for personal application achieved within the component; the extent to which patients can

attempt to apply the technique and problem-solve with therapist help when they encounter difficulties; and so on. Again, more is not necessarily better: for instance, there is good evidence that biofeedback does not enhance the learning or practice of relaxation skills;^{313,314} the results of Spence *et al* (1995) suggest that the use of biofeedback in teaching relaxation skills may undermine maintenance of those skills. Many examples of the variability of content within psychological therapies could be further listed. Work is in progress to explain why such differences occur, to quantify the effect of these differences upon overall outcome efficacy, and to give minimal standards for competent therapy. *Tables 36–38* demonstrate the diversity of treatment content and the plethora of measures, such that comparison and combination require more detailed examination of treatment. Early results in which adjustment for these differences has not been made are shown below, together with two studies which both achieved the highest quality score (see *Tables 34* and *35*) and which illustrate the extent of treatment gain possible.

Interim summary of averaged effect sizes over the studies by different outcome domains

Summary statistics for several analyses of effect sizes are presented in *Table 39*. The computations were performed using Stauffer's (1996) MetaQuik-16

*The following studies are excluded from the tables: Puder (1988), in which the comparison was of immediate versus delayed treatment; Strauss *et al* (1986), Pilowsky *et al* (1995), Peters *et al* (1991), and Cott *et al* (1992), because the treatments were either insufficiently detailed or too unorthodox to code satisfactorily.

TABLE 36 Treatment content: low back pain, mixed chronic pain or specific size studies

Study	Ed	Rel	Ex	Med	Beh	Goal	Att	Cog	Prob	Fam	Gen	F-up
Flor & Birbaumer 1993	-	(+)	-	=	-	-	+	+	-	-	-	12 m
Kerns et al 1986	CBT	CBT	-	B	B	=	CBT	CBT	B	-	-	6 m
Linton & Gotestam 1984	-	=	+	=	+	+	-	-	-	-	-	-
Linton et al 1985	-	+	-	-	+	+	-	-	-	-	-	1 w
Newton-John et al 1995	+	(+)	-	-	-	+	+	+	-	-	+	6 m
Nicholas et al 1991	C,CR	CR B,BR E	C,CR BR	B,BR E	B,BR	B,BR	C,CR B,BR	C,CR	-	-	-	12 m
Nicholas et al 1992	+	+	=	+	+	+	+	+	-	-	+	6 m
Turner 1982	-	=	-	-	-	+	+	+	-	-	-	18 m
Turner & Clancy 1988	-	CBT	B	-	=	-	CBT	CBT	-	B	-	12 m
Turner et al 1990	B,BE	-	E,BE	-	B,BE	B,BE	-	-	-	B,BE	-	12 m
Turner & Jensen 1993	-	R,CR	-	-	-	-	C,CR	C,CR	-	-	-	12 m
Vlaeyen et al 1995	-	RO	=	-	=	-	CO	CO	-	=	-	12 m
Altmaier et al 1992	=	+	-	-	-	-	-	+	+	+	-	6 m
Vlaeyen et al 1996	=	+	=	-	-	-	+	-	-	-	-	12 m
Spence et al 1995	-	=	-	-	-	-	+	-	-	-	-	6 m
Att	attention management		CR	cognitive and relaxation group			Prob	problem solving				
B	behaviour therapy group		E	exercise group			R	relaxation group				
BE	behaviour and exercise group		Ed	education			Rel	relaxation				
Beh	operant management		Ex	exercise programme			RO	relaxation and operant group				
BR	behaviour and relaxation group		Fam	social/family involvement			+	treatment group only				
C	cognitive group		F-up	follow-up period			(+)	different for treatment and control groups				
CBT	cognitive-behavioural therapy group		Gen	generalisation/maintenance			-	no intervention				
CO	cognitive and operant group		Goal	goal setting			=	both groups				
Cog	cognitive restructuring		Med	drug reduction								

TABLE 37 Treatment content: rheumatoid arthritis or osteoarthritis studies

Study	Ed	Rel	Ex	Med	Beh	Goal	Att	Cog	Prob	Fam	Gen	F-up
Keefe et al 1990a,b	(+)	+	(+)	-	+	+	+	+	-	-	?	6 m
Keefe et al 1996	ED	PC	-	-	PC	PC	PC	PC	-	=	-	-
Kraimaat et al 1995	-	CBT	OT	-	-	=	CBT	CBT	-	-	-	6 m
Parker et al 1988	+	+	-	-	-	=	+	-	+	-	-	12 m
O'Leary et al 1988	=	+	-	-	-	+	+	+	-	-	-	-
Bradley et al 1987	=	+	-	-	+	+	-	-	-	-	-	6 m
Appelbaum et al 1988	+	+	-	-	-	-	-	+	+	+	-	18 m
Radojevic et al 1982	ED	B	-	-	B	-	-	B	-	=	-	2 m
Att	attention management		Ex	exercise programme			PC	pain coping skills group				
B	behaviour therapy group		Fam	social/family involvement			Prob	problem solving				
Beh	operant management		F-up	follow-up period			Rel	relaxation				
CBT	cognitive-behavioural therapy group		Gen	generalisation/maintenance			+	treatment group only				
Cog	cognitive restructuring		Goal	goal setting			(+)	different for treatment and control groups				
ED	education group		Med	drug reduction			=	both groups				
Ed	education		OT	occupational therapy			-	no intervention				

TABLE 38 Comparison of treatment delivery

Study	Ed	Rel	Ex	Med	Beh	Goal	Att	Cog	Prob	Fam	Gen	F-up
Individual versus group treatment												
Spence <i>et al</i> 1989,1991	-	=	-	-	=	=	=	=	-	-	-	24 m
Inpatient versus outpatient treatment												
Williams <i>et al</i> 1996	=	=	=	=	=	=	=	=	-	=	=	12 m
Peters & Large 1990	IP	IP	IP	=	IP	OP	IP	IP	-	-	-	18 m
With versus without spouse/family involvement												
Moore & Chaney 1985	=	=	-	-	=	-	=	=	=	+	-	6 m
Keefe <i>et al</i> 1996	-	=	-	-	=	=	=	=	-	+	-	-
Radojevic <i>et al</i> 1982	-	=	-	-	+	-	-	=	-	+	-	2 m
Att	attention management		F-up	follow-up period			Prob	problem solving				
Beh	operant management		Gen	generalisation/maintenance			Rel	relaxation				
Cog	cognitive restructuring		Goal	goal setting			+	treatment group only				
Ed	education		IP	in-patient			-	no intervention				
Ex	exercise programme		Med	drug reduction			=	both groups				
Fam	social/family involvement		OP	out-patient								

TABLE 39 Effect sizes

Comparison group	Mean effect size	Number of treatment groups	95% CI	z
Domain 1: Pain experience – represents pain severity, intensity and affect. It is reflected in measures such as the McGill Pain Questionnaire ³²⁵ and VAS scores of intensity/affect.				
Waiting list control	0.29	30	0.14–0.43	3.91
Treatment as usual	0.27	23	0.11–0.43	3.3
Domain 2: Affect – as assessed by measures such as the Beck Depression Inventory, ²⁸⁶ West Haven–Yale Multi-dimensional Pain Inventory – affective distress, ³²⁶ Arthritis Impact Measurement – psychological disability, ³¹⁷ Spielberger State-Trait Anxiety Inventory, ³²⁹ Centre for Epidemiological Studies – depression. ³²⁸ Two analyses are reported.				
(a) Global mood measures – a variety of measures, including depression, are entered into the analysis.				
Waiting list control	0.32	30	0.15–0.49	3.8
Treatment as usual	0.12	21	-0.07–+0.32	1.25
(b) Depression measures – as estimated by standardised questionnaires.				
Waiting list control	0.30	27	0.13–0.47	3.37
Treatment as usual	-0.10	17	-0.34–+0.12	-0.92
Domain 3: Cognitive – measures of cognitive activity, mostly concerned with coping self statements as assessed by Cognitive Strategies Questionnaire, ²⁸⁴ Pain Beliefs Questionnaire, ³³⁰ Pain Cognitions Checklist. ³³¹ These are divided into two – positive coping measures and negative measures (catastrophising).				
(a) Positive – excluding catastrophising				
Waiting list control	0.37	17	0.21–0.52	4.77
Treatment as usual	0.36	17	0.24–0.49	5.86
(b) Catastrophising				
Waiting list control	0.71	10	0.46–0.96	5.53
Treatment as usual	0.47	12	0.29–0.65	5.14

continued

TABLE 39 contd Effect sizes

Comparison group	Mean effect size	Number of treatment groups	95% CI	z
Domain 4: Behaviour – assessed by a number of measures which fall into two broad categories, analysed separately: (a) active (normal) behaviour, e.g. walking distance, number of motor activities – West Haven–Yale Multidimensional Pain Inventory activity checklist; ³²⁶ (b) behavioural abnormality, e.g. Keefe & Block pain observation system. ³¹⁸				
(a) Increasing behavioural activity (normal)				
Waiting list control	0.44	11	0.21–0.68	3.8
Treatment as usual	0.22	4	0.08–0.36	3.16
(b) Abnormal pain behaviour				
Waiting list control	0.46	13	0.23–0.68	3.93
Treatment as usual	0.09	5	–0.8–+0.27	1.06
Domain 6: * Self-rated interference – measures which sample social role interference include Sickness Impact Profile – self-rated, ³²⁷ Arthritis Impact Measurement Scale (and Dutch equivalent, IRGL) – self-care, ³¹⁷ West Haven–Yale Multidimensional Pain Inventory – interference or general activity. ³²⁶ A further analysis could (will) be performed on ratings made by others (rather than self-rating).				
Waiting list control	0.43	25	0.25–0.60	4.81
Treatment as usual	0.19	14	0.00–0.39	2.02
*Domain 5: Biological measures (physiology and fitness) has not been included since the combination of different measures has not yet been resolved.				

program. The results are a first approximation but give a good enough estimate for present purposes. The general scheme of the analysis was to consider effect sizes for outcomes in various domains which represent conceptually distinct aspects of chronic pain. In order that the data are interpreted with caution, a number of the issues raised using this approach need to be taken into consideration.

1. Not all studies contributed measures to each domain.
2. Studies used different measures and the statistical relationship between these measures and the common constructs which conceptually underpin the domain is not known. The authors believe that there is enough commonality to justify the aggregation of different measures.
3. Some studies assesses the domain in more than one way. Where this has happened, one measure only has been selected to represent that domain. The rule employed was to select data from standardised multi-item assessments in preference to less psychometrically-sound measures.
4. Some studies contributed more than one effect size because they compared more than one

treatment with a control condition. These data have not been aggregated to estimate study effect sizes but the analysis has been maintained at the treatment group level. This may mean that the average effect sizes reported are biased.

5. Studies used various control groups. The present analysis compares treatment with two types of control – Waiting list (no treatment) and Treatment as usual. These analyses have not included trials in which an active treatment is compared with an educational programme.
6. Qualifiers for treatments were not incorporated into the analysis. These include methodological quality of trial, treatment content (see *Tables 36* and *37*), treatment delivery (see *Table 38*) initial severity of patients' problems, reliability of measures and length of follow-up.

While the effects of these modifications, once made, may increase or decrease effect sizes, the current status of the analysis (see *Table 39*) suggests that the effect sizes in most domains, although small, are reliably different from zero. This is of particular interest when the comparison control group is 'treatment as usual', rather than 'waiting list (no treatment)'.

Model studies

Keefe and colleagues, 1990a; b

Keefe and colleagues (1990a; b) at Duke University Medical Center, North Carolina, USA, compared pain coping skills training for patients with osteoarthritis of the knee with an arthritis education programme and with routine clinical care. The rationale arose, in part, from the limitations of pharmacotherapy and of surgery for a common painful and disabling condition, and, in part, from the discovery of relationships between cognitive coping strategies and disability in osteoarthritic patients which resembled that in other chronic pain populations. Arthritis education is widely used in the USA in a fairly standard format for this population,³¹⁵ and therefore constituted an appropriate comparison group.

A total of 99 diagnosed patients, mostly women (average age, 64 years; average pain duration, 12 years), were all given a patient-directed publication with information on arthritis and its treatment. Those in the arthritis education group also attended a group lecture (6–9 participants) and discussion which covered information on arthritis, treatments, exercises to maintain flexibility and strength, and joint protection and maintenance of mobility and function by correct posture and the selective use of aids. These sessions were for 90 minutes per week for 10 weeks.

Patients in the pain coping skills training group, described as CBT, met in the same size groups and for the same time as those in the arthritis education group. They received education concerning pain as a multidimensional experience involving thoughts, feelings and behaviour. Skills to enhance coping were taught: relaxation with imagery and distraction; activity–rest cycling and its application to returning to pleasant activities which had been abandoned; identifying and altering unhelpful thoughts and feelings associated with coping poorly with pain.

A questionnaire concerning expectations of treatment³¹⁶ indicated that the treatments were equally credible to patients in both intervention groups. Treatment integrity was measured by independent ratings of audiotapes from treatment, blind to their condition of origin, and was good. Attendance at at least nine of the ten sessions was achieved by over 80% of each group. Pre-treatment and post-treatment measures were taken of cognitive and behavioural coping strategies (CSQ³⁸⁴); of pain, physical disability and psychological

dysfunction (anxiety and depression) using a measure of demonstrated reliability and validity (AIMS³¹⁷); and of certain behaviours (such as guarding and rigidity) using an established observation method during a standardised set of activities.³¹⁸ There were no significant differences in pretreatment scores between groups, or in their drug intake (grouped into NSAIDs, narcotic analgesics, non-narcotic analgesics, and steroids). Age, sex and obesity index were co-varied in analysis of results.

At the end of treatment, patients in the CBT group scored significantly lower than both other groups on pain (about a 15% reduction of unadjusted scores) and psychological dysfunction (about a 25% reduction in unadjusted scores); those in the arthritis education group scored less on psychological dysfunction than those in routine care. There were no significant changes in drug intake or pain behaviours. For patients in the CBT group alone, reduction in physical disability was associated with improvement in cognitive and behavioural coping strategies used.

At 6-month follow-up, during which period patients had received three calls from their group therapist reviewing the use of pain coping skills for those in the skills training group, and general progress for those in the arthritis education group, the measures were repeated. Patients in the pain coping skills group had significantly lower physical disability scores than those in the arthritis education group, resulting from improvement over the follow-up period in the former group (particularly for women) and worsening in the latter, and lower psychological dysfunction scores than those in either of the other groups, although this did not represent significant improvement since the end of treatment. There was no significant difference in pain, pain behaviour, or in drug use.

Overall, this economical programme (in terms of time; staff, however, included very experienced psychologists to supervise therapists) produced substantial benefit in comparison with an education programme which required almost the same resources, and which, in its simple form without any substantial behavioural or cognitive component, failed to show any advantage over routine medical care. This has immediate practical implications (in the USA); the involvement of spouses, to improve generalisation and maintenance of skills, is now being tested; the addition of behavioural rehearsal of everyday tasks, using the taught coping methods, is under consideration.

Williams and colleagues, 1996

This study, by Williams and colleagues at St Thomas's Hospital, London, was initially funded by research monies but was subsequently adopted by the NHS. It ran from 1989 to 1992 and, at the time of planning, represented a rather modest form of inpatient and outpatient programme in the literature: since that time, however, the norms have become shorter and inpatient treatment is rarer.

Behavioural theory predicts that in-patients would have considerable problems in maintaining new healthier habits learned in the hospital environment when they returned to their home environments, so that the intensive programme would provide expensive short-term gain compared with more gradual but lasting gain for out-patients who established the new habits in their own environments. Nevertheless, given the rarity of outpatient programmes and the considerable travel difficulties encountered by most patients with chronic pain, even over short distances, inpatient programmes appeared to be the only form accessible to most patients.

In-patients in the study were self-caring in hostel accommodation attached to the project, with no staff available, except for emergencies, outside office hours; with some exceptions, they spent weekends at home. Out-patients attended for one afternoon per week for 8 weeks. Randomised patients were treated within the larger population of non-randomised patients: most patients preferred to opt for inpatient or outpatient treatment, the major determinant being the travelling distance – patients were referred from all over the UK. The mean age of patients was 50; mean duration of pain 8 years; 57% were female; less than 10% were employed, and about 60% received disability-related benefit as their main source of income.

A team of two psychologists, a part-time anaesthetist specialised in pain management, a physiotherapist, an occupational therapist and a nurse, all trained by the senior psychologist, delivered both inpatient and outpatient treatments, using the same methods and materials, and based on an explicit model of pain and problems secondary to pain.³¹⁹ Patients learned exercise and stretch to improve fitness, flexibility and strength and to remedy particular problems; they set activity goals to which exercise activities were related, and they learned to use activity–rest cycling to avoid over-activity (relative to their physical condition) or excessive rest. All exercises and activities increased on a quota system with explicit reinforcement of quota achievement. Drugs of no identifiable benefit were reduced

on a quota system. Relaxation was taught and practised in application to a range of everyday activities and sleep. Behavioural and cognitive teaching addressed the establishment of new habits, and the reduction of fears and depression by return to activity by graded increase (exposure) and by identifying and challenging overly-negative beliefs. Spouses and family members were encouraged to attend the latter part of both programmes.

Both in- and outpatient ratings of expectation of treatment benefit were similar. Post-treatment assessment (at 1, 6 and 12 months after the end of treatment) was carried out by raters blind to treatment condition, and included objective measures of physical performance.³²⁰ While there was minimal change in pain ratings, both in-patients and out-patients made significant improvement in psychological measures, physical performance, self-report of function, and drug intake, with inpatient superiority particularly in physical performance. At 1-year follow-up, with an overall gain in scores for both groups on most measures, inpatient superiority was evident in physical, cognitive and healthcare use (including drug intake) measures. These are expressed as NNTs in *Figure 28*.³²¹

In 1992, the inpatient programme cost just under £2000 per patient, and the outpatient programme just under £450. With recent changes in the NHS, the pain management programme (INPUT) now has its own budget: its running costs consist mainly of staff salaries and of patient accommodation, and (with overhead payments to the Hospital Trust) fees for in-patients are now £3750 and for out-patients, £1900. The average saving on drugs is £250 per patient per year;³²² other savings in healthcare costs are spread over several years and hard to estimate. Return to work is modest, although some non-working patients find their qualification for disability benefits reduced following treatment.

Conclusions

Research reports of the efficacy of psychological therapies for the treatment of chronic pain are legion. As with other pain research, reports of quality are rare. Of the 35 RCTs entered into the analysis there is considerable variability in the content and focus of therapy, the measures used and the target populations. Initial meta-analysis, with adjustment for sample size but for none of the study qualities described and discussed in this chapter, shows small effect sizes across domains

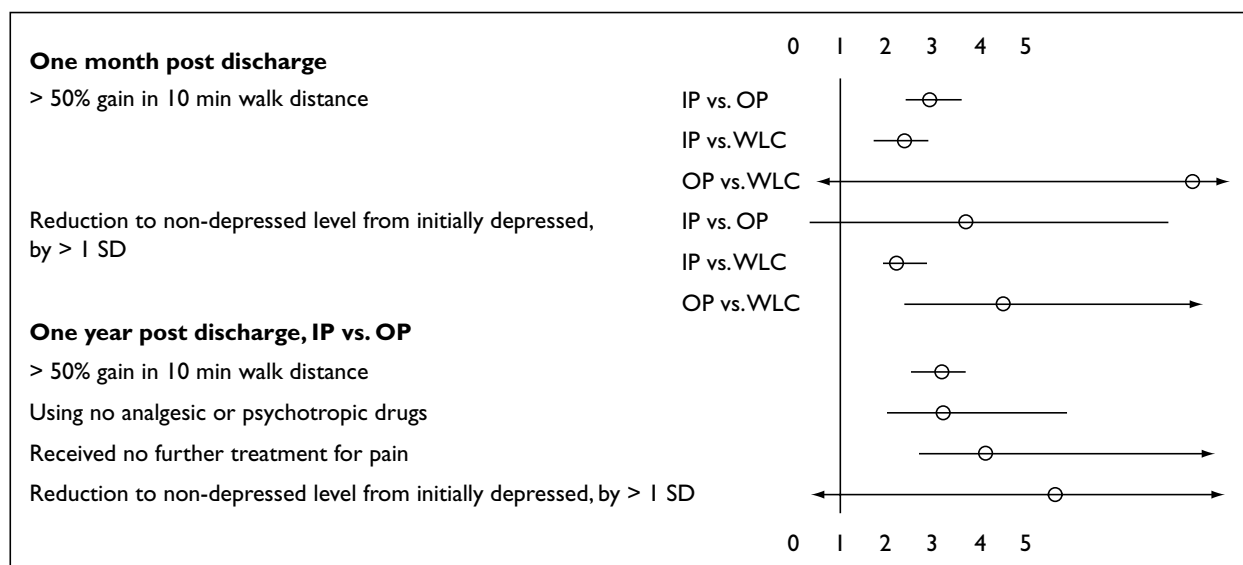


FIGURE 28 NNTs for in-patients, out-patients and waiting list controls (WLC) at 1 month and 1 year follow-up

of measurement (over a mean follow-up time of 9 months), of which most are significantly greater than zero. Further analysis adjusting for qualities of studies and grouping by treatment content is under way. The overall impression is of far less powerful treatment than suggested by the two model studies reviewed here. It is not yet certain that a more consistent picture will emerge from the detailed analyses planned.

Strengths are evident in this field that are transferable to other pain studies. Of particular note is the emphasis here on multiple outcome measures across a range of domains, process checks on blindness and the quality of therapists. Unique to this field is the importance of the accurate description of therapeutic components to enable meaningful analysis of outcome. Three related studies are in preparation.

1. A report on the best practice for the efficient use of electronic databases for psychological therapies. Medline is insufficient for these searches and a strategy for maximising available resources will be presented.
2. A conceptual analysis of the sources of error and variability in the quality of studies. Of particular concern is the extent of poor reporting of the theoretical basis of therapies, the importance of therapist training, the choice and use of measures. The recommendations of Altman³²³ are noted.
3. A meta-analysis of the 35 studies will provide the necessary substantive evidence. Preliminary evidence, reported above, of the high quality

trials demonstrates large and sustainable changes for targeted outcomes.

Recent systematic reviews of CBTs provide strong evidence for efficacy across a range of mental health problems.³¹¹ Much can be learned in pain research from the methodological sophistication of these studies. In addition to a summary statement on the efficacy of CBT in chronic pain, guidance on the necessary quality of both treatment content and research design will be offered.^{324,332}

Coding system

Source of information – origin according to search strategy, year of publication, journal of publication, country where study conducted, clinic status, academic affiliation of clinic, researcher's affiliation, patient residency during treatment.

Trial design – allocation/randomisation procedure, independence of randomisation from experimenters, exclusion and inclusion criteria, power calculations, treatment modes (i.e. individual, group, spouse/family session), length of treatment in weeks, hours of treatment, length of follow-up, therapist training, therapist supervision, level of therapist experience, therapist adherence to treatment protocol, assessment of therapist competence during trial, manualisation of treatment, patient rating of therapist effectiveness, patient pre-treatment rating of expectations of treatment, assessment of treatment credibility, confounding of therapists with treatment.

(continued)

Coding system *contd*

Participants – sampling strategy for population, source of sample, total sample size before selection or attrition, number entered into trial, number of subjects at end of treatment, number of subject, followed up, numbers of males and females, mean and SD of age, age range, mean chronicity of pain, mean number of previous non-surgical treatments, mean number of previous surgeries, percentage of sample with previous surgery, diagnostic label for sample, main site of pain, socio-economic status of sample, mean years of education, percentage married or cohabiting, percentage home-makers, percentage in paid employment, time since last worked for those not working, percentage with compensation claims outstanding, percentage on welfare benefit, source of income.

Chapter 20

Chronic pain services – evidence on costs

There is little information about the costs and benefits of chronic pain services, and what little there is barely constitutes evidence. Costs may be determined from the bottom up – contrasting, for instance, two or more different types of treatment for a condition and working out the costs and benefits for each. This method is precluded by lack of sufficient evidence; for instance, the fact that we have no real evidence that TENS works in chronic pain. Evidence of effectiveness must come first. Rational assessment of cost–benefit needs evidence of effectiveness.

The other way is to use a top-down approach, in which the disease burden is examined, changes are estimated, and judgement is made as to whether pain clinics add to costs or reduce them. Here, at least, there is some evidence but not very much and not very recent.

Prevalence

The prevalence of chronic non-malignant pain was estimated for the Oxford Region for the summer of 1982.³³³ The population served, 2.3 million, has

a Regional Pain Relief Unit with 1115 ‘actively maintained records’ of patients with non-malignant pain – records which had not been archived, excluding those who had died or not returned to the unit for 18 months.

This gives an overall prevalence of 485 patient per million population. However, the Unit treated patients from outside the region, and adjusting for that the prevalence would be lower, at 325 patients per million.

Referrals in 1982

Referral patterns for 1982 are shown in *Table 40*. Low back pain was the single most prevalent condition treated, although all neuralgias together constituted half the total, with 19% of cases comprising less than 2% of the total individually.

Changes since 1982

No documented evidence of change exists. Present patterns of referral and perceived changes include the following.

- Overall workloads have increased since 1982. Medical staffing has increased from one

TABLE 40 Referral patterns for the Oxford Regional Pain Relief Unit in 1982

Condition	Percentage of total	Mean age	Mean pain duration
Low back pain	26	52	9
PHN	11	73	3
Post-traumatic neuralgia	9	50	5
Atypical facial neuralgia	6	48	4
Intercostal neuralgia	5	55	4
TN	5	64	9
Perineal neuralgia	4	65	7
Abdominal neuralgia	4	56	6
Stump pain	3	63	13
Osteoarthritic hip	3	74	4
Sympathetic dystrophy (RSD)	2	59	4
Coccydynia	2	54	6
Cervical spondylosis	2	53	6
Other conditions	19		

consultant and senior registrar to two consultants, a University Reader in Pain Relief (part-time honorary consultant) and senior registrar.

- There are more specialist pain centres in the UK, and in the former Oxford Region there are consultants (especially anaesthetists) specialising in pain relief.
- More treatment now occurs in primary care, particularly in Oxfordshire. For example, postherpetic neuralgia will often now be treated by GPs with antidepressants.
- The types of patient referred have changed, tending now to be the more difficult cases, often with at least some attempt to treat having been made in primary care.³³⁴

Comment

It is likely that the burden of chronic pain as seen at pain clinics has increased substantially in the last decade. Patients are usually in the sixth and seventh decade of life, so demographic imperatives are likely to increase prevalence still further until about 2020.

Cost of chronic pain treatment

A detailed study of the costs incurred by users of speciality pain clinic services³³⁵ has shown that users of the services incur less direct healthcare expenditure than non-users with similar conditions.

Between January 1986 and April 1988, 626 patients were referred to the chronic pain clinic in Hamilton, Ontario, Canada; 210 did not attend the clinic ('non-attenders'), 180 had a consultation appointment only ('consultation only'), 98 did not complete the treatment programme ('incomplete treatment') and 83 completed the treatment programme. The use of different types of health services and other costs were calculated for a sample of 222 of the 626 patients, by asking

patients about their use of five categories of direct health services – primary care, emergency room and specialists, hospital episodes and days, and the use of seven types of other health professionals. Other direct and indirect costs for the patient associated with their use of health care, were also estimated. Results are given in the paper in 1991 dollars but it is not clear if these are Canadian or US dollars.

There was no demographic or condition-diagnosed difference between the four groups. The results showed that direct healthcare costs were lower for users of chronic pain services than for non-users. Costs for one year, broken down by type, are shown in *Table 41*.

The total annual direct health costs were much lower for users of chronic pain services (even if it was only a consultation), and the savings were clearly derived mostly from reduced costs of days spent in hospital. This is shown graphically in *Figure 29*.

The 74% of chronic pain referrals who actually used some chronic pain services used only 64% of the total costs for the referred patients (*Figure 30*). The 'saving' that resulted from using chronic pain services derived mainly from the intensive users of the service who had treatment, rather than those who had only a consultation.

Comment

These data are the clearest evidence available that chronic pain services not only benefit patients, but are also an efficient way of dealing with chronic pain in the community.

The average direct healthcare cost of a patient using chronic pain services, even if that was a single consultation, was \$2947. Referred patients who did

TABLE 41 Annual direct healthcare costs per patient, using different levels of service, referred to a chronic pain unit (1991 \$)

Service used	Non-attender	Consult only	Incomplete treatment	Complete treatment
Number surveyed	57	80	44	41
Primary care visits	477	422	412	462
Specialists	548	642	862	817
Emergency room	206	439	266	191
Hospital stay	3116	2017	462	1290
Health professional	833	396	226	237
Total	5181	3917	2229	2996

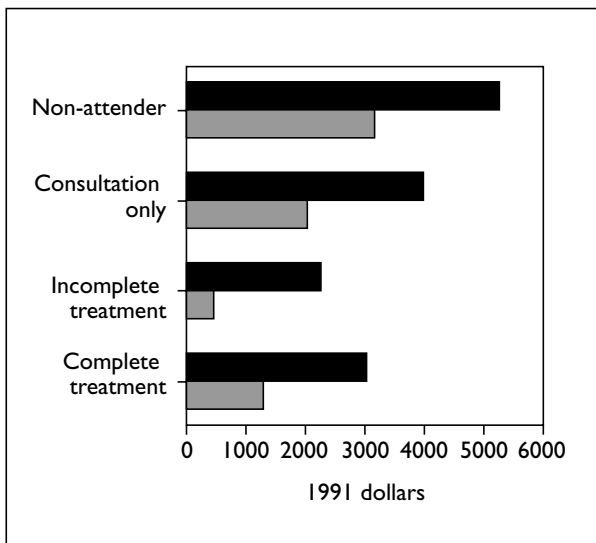


FIGURE 29 Total direct healthcare costs per patient (■) and cost of hospital stay (▣) for patients referred to a chronic pain unit using different levels of service (1991 \$)

not use the service cost more, on average \$5181. The difference between these averages was \$2234.

Using the most conservative estimate – that is, assuming no cost inflation since 1991 and that the currency was Canadian dollars (exchange rate approximately \$2 per £) – the average difference amounts to a saving in direct healthcare costs of about £1117 per patient. Using this figure in conjunction with the 1982 figure of 1115 patients

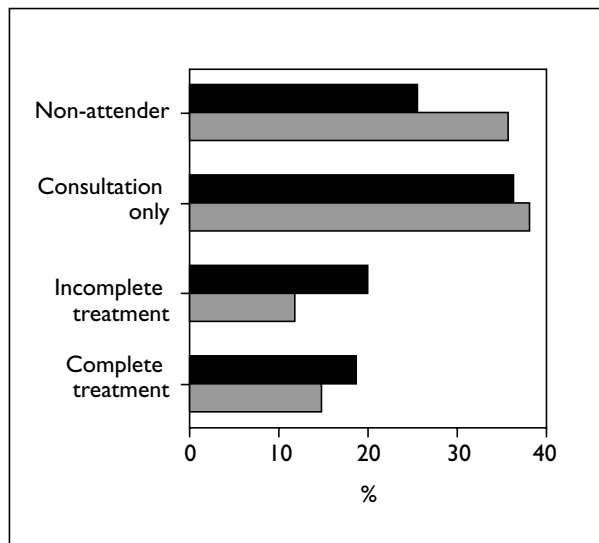


FIGURE 30 Percentage of total number (■) and total direct healthcare costs (▣) for patients referred to a chronic pain unit using different levels of service

with non-malignant pain on the Oxford Regional Pain Relief Unit books translates to a healthcare saving of £1,250,000.

This compares with the present (1996) running cost (labour, consumables, estates, overheads) of the Oxford Pain Relief Unit (with a larger workload) of £600,000. Put simply, the evidence indicates that the use of pain clinics results in direct health service savings equal to twice their cost.

Chapter 21

Conclusions

Determining whether a service is worthwhile involves a number of different issues. It involves knowing whether the various components of the service (interventions) are effective, how much they cost, and examining whether their delivery is efficient. This review has concentrated on trying to determine whether interventions in chronic pain can be shown to be effective through systematic review. The report is not comprehensive, because the Health Technology Assessment grant was for 1 year only.

Strategies for finding and assessing the quality of relevant studies and reviews have been developed. Over 150 systematic reviews relating to chronic pain have been found. A system of quality scoring the reviews was applied which showed that high quality reviews are significantly less likely to give a positive result than reviews of lower quality. The simple quality-scoring system for RCTs that was devised has wider applications than just pain studies.

Many, if not most, of the interventions commonly used in chronic pain treatment, and covered in this review, can be shown to be very effective. Some can be shown to be ineffective. While these findings support much of current practice in chronic pain treatment, a common theme is that knowledge remains limited. In particular, information on which to base economic analyses is missing. However, such information as is available indicates that pain clinics result in direct healthcare savings of over £1000 per patient per year, and that total savings may be twice the cost of the chronic pain service. Since major demographic changes will affect the NHS over the next few decades, and ageing populations will demand more chronic (and cancer) pain therapy, the provision of more information on economic as well as humanitarian benefits will be important.

Our main conclusion is that recommendations can be made with confidence about the direction of future research in chronic pain, to improve knowledge of the most effective treatments, to determine which patients may get maximum benefit from particular treatments, and to initiate research into the most efficient type and delivery of service.

Effective interventions (of those covered)

- Minor analgesics.
- Anticonvulsant drugs.
- Antidepressant drugs.
- Systemic local anaesthetic-type drugs.
- Topical NSAIDs in rheumatological conditions.
- Topical capsaicin in diabetic neuropathy.
- Epidural corticosteroids for back pain and sciatica.
- Psychological interventions.

Interventions where evidence is lacking (of those covered)

- Transcutaneous electrical nerve stimulation (TENS) in chronic pain.
- Relaxation.
- Spinal cord stimulation (SCS).

Ineffective interventions (of those covered)

- Intravenous regional sympathetic blockade (IRSB).
- Injections of corticosteroids in or around shoulder joints for shoulder pain.

Costs

The information available on costs and benefits of chronic pain services in the UK is limited; however, there is evidence that the use of pain clinics results in direct health service savings equal to twice their cost.

Comment

First, reviews are necessary of those interventions which we, for reasons of time, were unable to cover.

Second, it is clear that high quality RCTs are needed in a number of different areas covered by this review. We simply do not know whether TENS, a widely-used intervention in chronic pain (used over 500,000 times per year in Canada¹¹⁵), works. Also, even where there is evidence, it is

impossible to say, for instance, which is the best antidepressant, or best anticonvulsant, or whether the best antidepressant is better than the best anticonvulsant.

Such trials have conventionally used relatively small numbers of patients per group – of the order of 50 or less. These may be enough to demonstrate a statistical difference where effects differ between groups by 30% or more, and are typical of studies carried out with industrial support to establish effectiveness for registration or licensing purposes, but such trials cannot establish the ‘true’ clinical relevance of an intervention. On-going work for a subsequent report shows that, for the same difference in effect between groups, the size of a trial to determine the ‘true’ NNT (plus or minus 0.5 units) would need to be of the order of 500 patients per group.

These trials require such large numbers of patients that they are beyond the capability of any one centre to run. This approach dictates multicentre studies with a common protocol and central randomisation, which would need to be not only to be UK-wide but possibly Europe-wide to provide the requisite level of evidence. A further likely benefit from such large trials would be the organisation of single-patient data collection. Logistic analysis could then be used to determine whether particular patient characteristics affected successful or unsuccessful outcome, or particular method of treatment. Such trials need not be exorbitantly expensive. Following the model of the ‘mega’ cardiac trials, it is simple outcomes on effectiveness and safety that need to be gathered but for large numbers of patients.

Third, chronic pain services, like many areas of medicine, need process analysis. The evidence for efficacy of particular interventions in this review could be used to study the way in which the process of chronic pain is delivered. Patients with pain from multiple sclerosis need help. Even if there is a paucity of interventions with proven efficacy, they still need help. It is this aspect of

chronic pain care, which makes patients feel better, that is lost in focusing exclusively on intervention efficacy. Similarly, attempts to address cost-effectiveness by making single interventions are naive. As the cost analysis in chapter 20 shows, there is benefit to the NHS as a whole if the patients’ needs are met in chronic pain services, rather than in continued resort to multiple services.

Research recommendations

To tackle the lack of information about intervention efficacy

- Establish a single UK centre to organise large multicentre studies into the clinical relevance of existing and new chronic pain therapies.
- Encourage Europe-wide cooperation through the International Association for the Study of Pain and European special interest groups and health technology directorates.
- Provide strong advisory backup in trial design and ethics.
- Provide single-source trial monitoring to ensure quality.
- Provide single-source statistical and mathematical back-up.
- Provide a single centre for randomisation.

To address issues of cost-effectiveness

Chronic pain services provide an opportunity to explore the technique of process analysis applied to a chronic service. Snap-shot cost-effectiveness studies are of limited benefit in a chronic and complicated setting. The efficacy data provided here could be used to study the service as a whole. The methods could be of great value to other chronic healthcare services.



Acknowledgements

This report includes much work contributed by others in the course of research programmes in Oxford over the past 6 years. Alex Jadad, now in Hamilton, Ontario, Canada, and our team of volunteers (one of whom, Ron Martin, has been working with us for nearly 6 years) broke the back of obtaining nearly 15,000 RCTs in pain; these have been donated to the Cochrane Collaboration to appear in future versions of the Cochrane Library. Alex Jadad, together with Dawn Carroll and Phil Wiffen, did much to begin the task of sifting through the vast quantity of literature to extract the evidence on treatments for pain relief. David Gavaghan, Oxford University Computing

Laboratory, modelled the pain studies. Many others have been of direct help, in particular, Martin Tramèr, Eija Kalso (University of Helsinki), Sally Collins, Trish Green, Bethany Nye, Kate Seers, Göran Leijon (Linköping, Sweden) and Owen Moore; our thanks are also due to our librarians, Claire Abbott and Jo Riordan.

Many others have helped by commenting, challenging and, sometimes, changing our ideas as they evolved. Our thanks are due to, among others, Muir Gray, Dave Sackett, Iain Chalmers, Jon Deeks, Chris Glynn, Tim Jack, Douglas Justins, Peter Evans, Ed Charlton and Geoff Gourlay.



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**:63S–64S. [10951]*
2. Chalmers I, Dickersin K, Chalmers TC. Getting to grips with Archie Cochrane's agenda. *BMJ* 1992;**305**:786–7. [7693]
3. Jadad AR, McQuay HJ. A high-yield strategy to identify randomized controlled trials for systematic reviews. *Online J Curr Clin Trials* 1993; 27 Feb; Doc No 3. [7949]
4. Simes RJ. Confronting publication bias: a cohort design for meta-analysis. *Stat Med* 1987;**6**:11–29. [10157]
5. Hetherington J, Dickerson K, Chalmers I, Meinert CL. Retrospective and prospective identification of unpublished controlled trials: lessons from a survey of obstetricians and pediatricians. *Pediatrics* 1989;**84**:374–80. [6990]
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. [12326]
7. Jadad AR, Moore RA, Carroll D, *et al.* Assessing the quality of reports of randomized clinical trials: is blinding necessary? *Control Clin Trials* 1996;**17**:1–12. [9292]
8. Reiffenstein RJ, Schiltroth AJ, Todd DM. Current standards in reported drug trials. *Can Med Assoc J* 1968;**99**:1134–5. [10147]
9. Chalmers TC, Smith H, Blackburn B, *et al.* A method for assessing the quality of a randomized control trial. *Control Clin Trials* 1981;**2**:31–49. [8129]
10. Tyson JE, Furzan JA, Reisch JS, Mize SG. An evaluation of the quality of therapeutic studies in perinatal medicine. *J Pediatr* 1983;**102**:10–13. [10144]
11. Andrew A. Method for assessing of the reporting standard of clinical trials with roentgen contrast media. *Acta Radiol Diag* 1984;**25**:55–8. [10309]
12. 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. [10396]
13. Poynard T. Evaluation de la qualité méthodologique des essais thérapeutiques randomisés. *Presse Med* 1988;**17**:315–8. [10487]
14. 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. [10188]
15. Gøtzsche PC. Methodology and overt and hidden bias in reports of 196 double blind trials of non-steroidal antiinflammatory drugs in rheumatoid arthritis. *Control Clin Trials* 1989;**10**:31–56. (erratum: *Control Clin Trials* 1989;**50**:356) [10311]
16. Reisch JS, Tyson JE, Mize SG. Aid to the evaluation of therapeutic studies. *Pediatrics* 1989;**84**:815–27. [10310]
17. ter Riet G, Kleijnen J, Knipschild P. Acupuncture and chronic pain: a criteria-based meta-analysis. *J Clin Epidemiol* 1990;**43**:1191–9. [6010]
18. Brown SA. Measurement of quality of primary studies for meta-analysis. *Nurs Res* 1991;**40**:352–5. [10308]
19. Koes BW, Bouter LM, Beckerman H, van der Heijden GJ, Knipschild PG. Physiotherapy exercises and back pain: a blinded review [see comments]. *BMJ* 1991;**302**:1572–6. [4465]
20. Koes BW, Assendelft WJ, van der Heijden GJ, Bouter LM, Knipschild PG. Spinal manipulation and mobilisation for back and neck pain: a blinded review [see comments]. *BMJ* 1991;**303**:1298–303. [4875]
21. 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. [8119]
22. Nurmohamed MT, Rosendaal FR, Buller HR, *et al.* Low molecular-weight heparin versus standard heparin in general and orthopaedic surgery: a meta-analysis. *Lancet* 1992;**340**:152–6. [10305]
23. Oxman AD, Guyatt GH. Validation of an index of the quality of review articles. *J Clin Epidemiol* 1991;**44**:1271–8. [10196]
24. Fink A, Kosecoff J, Chassin M, Brook RH. Consensus methods: characteristics and guidelines for use. *Am J Public Health* 1984;**74**:979–83. [10486]
25. Feinstein AR. *Clinimetrics*. Newhaven: Yale University Press; 1987. [10484]

* Numbers in square brackets refer to the authors' database.

26. 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. [3935]
27. 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. [10187]
28. Sackett DL, Gent M. Controversy in counting and attributing events in clinical trials. *N Engl J Med* 1979;**301**:1410–2. [10405]
29. Peto R, Pike MC, Armitage P, *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. [8451]
30. 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. [17950]
31. Keele KD. The pain chart. *Lancet* 1948;**ii**:6–8. [609]
32. Scott J, Huskisson EC. Graphic representation of pain. *Pain* 1976;**2**:175–84. [600]
33. Wallenstein SL, Heidrich III G, Kaiko R, Houde RW. Clinical evaluation of mild analgesics: the measurement of clinical pain. *Br J Clin Pharmacol* 1980;**10**:319S–27S. [273]
34. Littman GS, Walker BR, Schneider BE. Reassessment of verbal and visual analogue ratings in analgesic studies. *Clin Pharmacol Ther* 1985;**38**:16–23. [9478]
35. 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. [10667]
36. Murphy DE, 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. [1374]
37. 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. [7410]
38. Jones B, Jarvis P, Lewis JA, Ebbutt AF. Trials to assess equivalence: the importance of rigorous methods. *BMJ* 1996;**313**:36–9. [18991]
39. McQuay H, Moore A. Placebos are essential when extent and variability of placebo response are unknown. *BMJ* 1996;**313**:1008. [1002]
40. Wall PD. The placebo and the placebo response. In: Wall PD, Melzack R, editors. *Textbook of pain*. Edinburgh: Churchill-Livingstone, 1994:1297–308. [10942]
41. Wall PD. The placebo effect: an unpopular topic. *Pain* 1992;**51**:1–3. [10626]
42. Beecher HK. The powerful placebo. *JAMA* 1955;**159**:1602–6. [10560]
43. Evans FJ. The placebo response in pain reduction. *Adv Neurol* 1974;**4**:289–96. [9388]
44. Porter EJB, Rolfe M, McQuay HJ, Moore RA, Bullingham RES. Single dose comparison of bicipadine, codeine and placebo in postoperative pain. *Curr Ther Res* 1981;**30**:156–60. [185]
45. 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. [65]
46. McQuay HJ, Bullingham RE, Moore RA, *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. [1037]
47. McQuay HJ, Carroll D, Poppleton P, Summerfield RJ, Moore RA. Fluradoline and aspirin for orthopedic postoperative pain. *Clin Pharmacol Ther* 1987;**41**:531–6. [824]
48. 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. [1675]
49. 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. [5665]
50. Gracely RH, Dubner R, Deeter WR, Wolskee PJ. Clinicians’ expectations influence placebo analgesia. *Lancet* 1985;**i**:43. [10456]
51. Wall PD. Pain and the placebo response. In: *Experimental and theoretical studies of consciousness*. CIBA Foundation Symposium 174. Chichester: Wiley, 1993:187–216. [9398]
52. Bergmann J-F, Chassany O, Gandiol 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. [11476]
53. 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. [11486]
54. Gøtzsche PC. Meta-analysis of NSAIDs: contribution of drugs, doses, trial designs, and meta-analytic techniques. *Scand J Rheumatol* 1993;**22**:255–60. [10488]

55. Moertel CG, Ahmann DL, Taylor WF, Schwartz N. Relief of pain by oral medications. *JAMA* 1974;**229**:55–9. [704]
56. Jadad AR. Meta-analysis of randomised clinical trials in pain relief [DPhil thesis]. Oxford: University of Oxford, 1994. [11299]
57. Moore A, McQuay H, Gavaghan D. Deriving dichotomous outcome measures from continuous data in randomised controlled trials of analgesics. *Pain* 1996;**66**:229–37. [11925]
58. 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. [29]
59. 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. [8700]
60. 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. [823]
61. 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. [333]
62. 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]. [1076]
63. 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. [4028]
64. 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. [8099]
65. 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. [11905]
66. 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. [11825]
67. Cook RJ, Sackett DL. The number needed to treat: a clinically useful measure of treatment effect. *BMJ* 1995;**310**:452–4. [11525]
68. Dudewicz EJ, Mishra SM. Modern mathematical statistics. New York: Wiley, 1988. [11895]
69. Moore RA, McQuay HJ, Gavaghan DJ. Deriving dichotomous outcome measures from continuous data in randomised controlled trials of analgesics: verification from independent data. *Pain* 1997;**69**:127–30. [18100]
70. Silverstein FE, Graham DY, Senior JR, *et al*. Misoprostol reduces serious gastrointestinal complications in patients with rheumatoid arthritis receiving nonsteroidal anti-inflammatory drugs. *Ann Intern Med* 1995;**123**:241–9. [17530]
71. Oxman AD, Guyatt GH. Guidelines for reading literature reviews. *Can Med Assoc J* 1988;**138**:697–703. [10169]
72. Vander Stichele RH, Dezeure EM, Bogaert MG. Systematic review of clinical efficacy of topical treatments for head lice. *BMJ* 1995;**311**:604–8. [18890]
73. 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. [12306]
74. L'Abbé KA, Detsky AS, O'Rourke K. Meta-analysis in clinical research. *Ann Intern Med* 1987;**107**:224–33. [11009]
75. Zhang WY, Li WanPo A. The effectiveness of topically applied capsaicin. A meta-analysis. *Eur J Clin Pharmacol* 1994;**46**:517–22. [12610]
76. McQuay H, Carroll D, Jadad AR, Wiffen P, Moore A. Anticonvulsant drugs for management of pain: a systematic review. *BMJ* 1995;**311**:1047–52. [11885]
77. McQuay H, Nye BA, Carroll D, Wiffen PJ, Tramèr M, Moore RA. A systematic review of antidepressants in neuropathic pain. *Pain* 1996;**68**:217–27. [13305]
78. 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. [11915]
79. Soll JC, McQueen MC. Respiratory distress syndrome. In: Sinclair JC, Bracken ME, editors. Effective care of the newborn infant. Oxford: Oxford University Press, 1992:333. [18860]
80. 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. Oxford: Health Technology Evaluation Association, 1995. (<http://www.jr2.ox.ac.uk/Bandolier/bandopubs/hpyl/hp0.html>). [18870]
81. 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. [7906]

82. Sinclair JC, Bracken MB. Clinically useful measures of effect in binary analyses of randomized trials. *J Clin Epidemiol* 1994;**47**:881–9. [18480]
83. 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. In preparation 1997. [18150]
84. Sackett DL, Deeks JJ, Altman DG. Down with odds ratios! *Evidence-based Med* 1996;**1**:164–6. [2006]
85. Deeks J. What the heck's an odds ratio? *Bandolier* 1996;**3** no 3:6–7. [1681]
86. Cochrane Collaboration Review Manager Software (RevMan). 1996. [18880]
87. Glass GV. Primary, secondary, and meta-analysis of research. *Educ Res* 1976;**5**:3–8. [10238]
88. Moore RA, McQuay H, Gray M. *Bandolier – the first 20 issues*. Oxford: Bandolier, 1995. [18910]
89. Sackett D, Rosenberg W, Haynes B. *Evidence based medicine*. London: Churchill Livingstone, 1996. [1303]
90. Moore TJ. *Deadly medicine*. New York: Simon & Schuster, 1995. [13285]
91. Eyspasch E, Lefering R, Kum CK, Troidl H. Probability of adverse events that have not yet occurred: a statistical reminder. *BMJ* 1995;**311**:619–20. [13275]
92. Jadad AR, McQuay HJ. Meta-analyses to evaluate analgesic interventions: a systematic qualitative review of their methodology. *J Clin Epidemiol* 1996;**49**:235–43. [10947]
93. 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. [10137]
94. Gerbarg ZB, Horwitz RI. Resolving conflicting clinical trials: Guidelines for meta-analysis. *J Clin Epidemiol* 1988;**41**:503–9. [10217]
95. Kassirer JP. Clinical trials and meta-analysis: What do they do for us? *N Engl J Med* 1992;**327**:273–4. [10206]
96. Oxman AD, Guyatt GH, Singer J, *et al*. Agreement among reviewers of review articles. *J Clin Epidemiol* 1991;**44**:91–8. [10148]
97. Patel M, Gutzwiller F, Paccaud F, Marazzi A. A meta-analysis of acupuncture for chronic pain. *Int J Epidemiol* 1989;**18**:900–6. [10361]
98. Chalmers TC, Berrier J, Hewitt P, *et al*. Meta-analysis of randomized controlled trials as a method of estimating rare complications of non-steroidal anti-inflammatory drug therapy. *Aliment Pharmacol Ther* 1988;**2** suppl 1:9–26. [10350]
99. 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. [10368]
100. 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. [10337]
101. Shekelle PG, Adams AH, Chassin MR, Hurwitz EL, Brook RH. Spinal manipulation for low-back pain [see comments]. *Ann Intern Med* 1992;**117**:590–8. [10117]
102. Ottenbacher K, DiFabio RP. Efficacy of spinal manipulation/mobilization therapy. A meta analysis. *Spine* 1985;**10**:833–7. [1116]
103. Anderson R, Meeker WC, Wirick BE, Mootz RD, Kirk DH, Adams A. A meta-analysis of clinical trials of spinal manipulation [see comments]. *J Manipulative Physiol Ther* 1992;**15**:181–94. [10377]
104. 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. [10541]
105. Felson DT, Anderson JJ, Meenan RF. The comparative efficacy and toxicity of second-line drugs in rheumatoid arthritis. *Arthritis Rheum* 1990;**33**:1449–61. [10564]
106. Felson DT, Anderson JJ, Meenan RF. Use of short-term efficacy/toxicity tradeoffs to select second-line drugs in rheumatoid arthritis. *Arthritis Rheum* 1992;**35**:1117–25. [10577]
107. 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. [10316]
108. 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. [10557]
109. Gam AN, Thorsen H, Lonnberg F. The effect of low-level laser therapy on musculoskeletal pain: a meta-analysis. *Pain* 1993;**52**:63–6. [10385]
110. Schmader KE, 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. [10356]
111. Lycka BA. Postherpetic neuralgia and systemic corticosteroid therapy. Efficacy and safety. *Int J Dermatol* 1990;**29**:523–7. [7927]
112. Naldi L, Zucchi A, Brevi A, Cavaliere D'Oro L, Cainelli T. Corticosteroids and post-herpetic neuralgia. *Lancet* 1990;**336**:947. [10455]

113. 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. [10374]
114. Melzack R, Wall PD. Pain mechanisms: a new theory. *Science* 1965;**150**:971–8. [19005]
115. Reeve J, Menon D, Corabian P. Transcutaneous electrical nerve stimulation (TENS): a technology assessment. *Int J Technol Assess Health Care* 1996;**12**:299–324. [1530]
116. 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, U.S. Department of Health and Human Services. AHCPR Publication No. 92-0032. Rockville, MD: Public Health Service; Feb 1992:24–5. [12186]
117. Royal College of Surgeons and College of Anaesthetists. Report of the working party on pain after surgery. London: Royal College of Surgeons, 1990. [12206]
118. McCaffery M, Beebe A. Pain: a clinical manual for nursing practice. St. Louis, Missouri: C.V. Mosby Company; 1989. [1672]
119. 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. [1670]
120. 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. [1676]
121. Park G, Fulton B. The management of acute pain. Oxford: Oxford University Press, 1991. [1679]
122. Woolf CJ, Thomson JW. Stimulation-induced analgesia: transcutaneous electrical nerve stimulation (TENS) and vibration. In: Wall PD, Melzack R, editors. Textbook of pain. Edinburgh: Churchill Livingstone, 1994:1191–217. [1671]
123. Justins DM, Richardson PH. Clinical management of acute pain. *Br Med Bull* 1991;**47**:561–83. [1528]
124. Carroll D, Tramer T, McQuay H, Nye B, Moore A. Transcutaneous electrical nerve stimulation in labour pain: a systematic review. *Br J Obstet Gynaecol* 1997;**104**:169–75. [13780]
125. 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. [1666]
126. Warfield CA, Stein JM, Frank HA. The effect of transcutaneous electrical nerve stimulation on pain after thoracotomy. *Ann Thorac Surg* 1985;**39**:462–5. [1680]
127. Pike PM. Transcutaneous electrical stimulation. Its use in the management of postoperative pain. *Anaesthesia* 1978;**33**:165–71. [1673]
128. VanderArk GD, McGrath KA. Transcutaneous electrical stimulation in treatment of postoperative pain. *Am J Surg* 1975;**130**:130–8. [1678]
129. 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. [1663]
130. 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. [1660]
131. 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. [1661]
132. 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. [1668]
133. 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. [1659]
134. Steptoe P, Bo JO. Transkutan nervestimulationer smertelindrende effekt ved fødsler. *Ugeskr Laeger* 1984;**146**:3186–9. [1662]
135. Wattrisse G, Leroy B, Dufosse F, Bui Huu Tai R. Electrostimulation cerebrale transcutanee: Étude comparative des effets de son association a l'anesthésie peridurale par bupivacaine-fentanyl au cours de l'analgésie obstetricale. *Cah Anesthesiol* 1993;**41**:489–95. [1664]
136. 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. [1667]
137. 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. [1677]
138. Rooney SM, Jain S, Goldiner PL. Effect of transcutaneous nerve stimulation on postoperative pain after thoracotomy. *Anesth Analg* 1983;**62**:1010–2. [1674]
139. McQuay H, Moore A. Epidural anaesthesia and low back pain after delivery [letter]. *BMJ* 1996;**312**:581. [13770]

140. Johnson MI, Ashton CH, Thompson JW. Long term use of transcutaneous electrical nerve stimulation at Newcastle Pain Relief Clinic. *J R Soc Med* 1992;**85**:267–8. [6770]
141. Jadad AR, Carroll D, Glynn CJ, McQuay HJ. 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* 1995;**10**:13–20. [11400]
142. Schutzer SF, Gossling HR. The treatment of reflex sympathetic dystrophy syndrome. *J Bone Joint Surg* 1984;**66A**:625–9. [8368]
143. Schwartzman RJ, McLellan TL. Reflex sympathetic dystrophy: a review. *Arch Neurol* 1987;**44**:555–61. [8369]
144. Procacci P, Maresca M. Reflex sympathetic dystrophies and algodystrophies: historical and pathogenic considerations. *Pain* 1987;**31**:137–46. [8366]
145. Hannington-Kiff JG. Intravenous regional sympathetic block with guanethidine. *Lancet* 1974;**i**:1019–20. [8358]
146. Loh L, Nathan PW, Schott GD. Painful peripheral states and sympathetic blocks. *J Neurol Neurosurg Psychiatry* 1978;**41**:664–71. [8364]
147. Hannington-Kiff JG. Pharmacological target blocks in painful dystrophic limbs. In: Wall PD, Melzack R, editors. Textbook of pain. London: Churchill Livingstone, 1989:754–66. [1565]
148. Bonelli S, Conoscente F, Movilia PG, Restelli L, Francucci B, Grossi E. Regional intravenous guanethidine vs stellate ganglion block in reflex sympathetic dystrophies: a randomized trial. *Pain* 1983;**16**:297–307. [8351]
149. Kettler RE, Abram SE. Intravenous regional droperidol in the management of reflex sympathetic dystrophy: a double-blind placebo-controlled, crossover study. *Anesthesiology* 1988;**69**:933–6. [8362]
150. Hanna MH, Peat SJ. Ketanserin in reflex sympathetic dystrophy. A double-blind placebo controlled trial. *Pain* 1989;**38**:145–50. [8357]
151. Rocco AG, Kaul AF, Reisman RM, Gallo JP, Lief PA. A comparison of regional intravenous guanethidine and reserpine in reflex sympathetic dystrophy: a controlled, randomized, double-blind crossover study. *Clin J Pain* 1989;**5**:205–9. [8367]
152. Blanchard J, Ramamurthy S, Walsh N, Hoffman J, Schoenfeld L. Intravenous regional sympatholysis: a double-blind comparison of guanethidine, reserpine, and normal saline. *J Pain Symptom Manag* 1990;**5**:357–61. [8350]
153. Hord AH, Rooks MD, Stephens BO, Fleming LL. Intravenous regional bretylium for the treatment of reflex sympathetic dystrophy. *Anesth Analg* 1992;**74**:818–21. [8360]
154. Dhar S, McGlone F, Dean J, Klenerman L. The temporal significance of the placebo response in reflex sympathetic dystrophy. Proceedings of the British Orthopaedic Research Society Meeting, 1992; London. 16–17. [8354]
155. Levine JD, Fye K, Heller P, Basbaum AI, Whiting-O'Keefe Q. Clinical response to regional intravenous guanethidine in patients with rheumatoid arthritis. *J Rheumatol* 1986;**13**:1040–43. [8363]
156. Sharpe E, Milaszkiwicz R, Carli F. A case of prolonged hypotension following intravenous guanethidine block. *Anaesthesia* 1987;**42**:1081–4. [8370]
157. Casale R, Glynn C, Buonocore M. The role of ischaemia in the analgesia which follows Bier's block technique. *Pain* 1992;**50**:169–75. [8352]
158. Koes BW, Scholten RPM, Mens JMA, Bouter LM. Efficacy of epidural steroid injections for low-back pain and sciatica: a systematic review of randomized clinical trials. *Pain* 1995;**63**:279–88. [13750]
159. Watts RW, Silagy CA. A meta-analysis on the efficacy of epidural corticosteroids in the treatment of sciatica. *Anaesth Intensive Care* 1995;**23**:564–9. [12725]
160. Kane RE. Neurologic deficits following epidural or spinal anesthesia. *Anesth Analg* 1981;**60**:150–161. [9334]
161. Turner JA, Loeser JD, Bell KG. Spinal cord stimulation for chronic low back pain; a systematic literature synthesis. *Neurosurgery* 1995;**37**:1088–96. [1546]
162. van der Heijden CJMG, van der Windt DAWM, Kleijnen J, Koes BW, Bouter LM. Steroid injections for shoulder disorders: a systematic review of randomized clinical trials. *Br J Gen Pract* 1996;**46**:309–16. [19040]
163. Jacox A, Carr DB, Payne R. Management of cancer pain. Clinical Practice Guideline No. 9. US Agency for Health Care Policy and Research. Rockville, MD: Public Health Service; 1994 March. AHCPR Publication No. 94-0592:41–74. [11034]
164. Rappaport ZH, Devor M. Trigeminal neuralgia: the role of self-sustaining discharge in the trigeminal ganglion. *Pain* 1994;**56**:127–38. [11202]
165. Loeser JD. Tic douloureux and atypical facial pain. In: Wall PD, Melzack R, editors. Textbook of pain. London: Churchill Livingstone, 1994:699–710. [11172]
166. Twycross R. The management of pain in cancer. In: Nimmo WS, Rowbotham DJ, Smith G, editors. Anaesthesia. Oxford: Blackwell Scientific Publications, 1994:1635–51. [1651]
167. Monks R. Psychotropic drugs. In: Wall PD, Melzack R, editors. Textbook of pain. London: Churchill Livingstone, 1994:963–89. [11182]

168. Reynolds JEF. Martindale: the extra pharmacopoeia. London: Pharmaceutical Press, 1993:292–314. [1643]
169. Grahame-Smith DG, Aronson JK. Oxford textbook of clinical pharmacology and drug therapy. Oxford: Oxford University Press, 1992:433–6, 443–4, 529–717. [1330]
170. Rall TW, Schleifer LS. Drugs effective in the therapy of the epilepsies. In: Goodman Gilman A, Rall TW, Nies AS, Taylor P, editors. The pharmacological basis of therapeutics. Toronto: McGraw-Hill, 1992:436–62. [1641]
171. 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. [8371]
172. Kienast HW, Boshes LD. Clinical trials of carbamazepine in suppressing pain. *Proc Inst Med Chicago* 1968b;**27**:50. [1593]
173. Keczek K, Basheer AM. Do corticosteroids prevent post-herpetic neuralgia? *Br J Dermatol* 1980;**102**:551–5. [1590]
174. Campbell FG, Graham JG, Zilkha KJ. Clinical trial of carbamazepine (Tegretol) in trigeminal neuralgia. *J Neurol Neurosurg Psychiatry* 1966;**29**:265–7. [1309]
175. Killian JM, Fromm GH. Carbamazepine in the treatment of neuralgia. Use of side effects. *Arch Neurol* 1968;**19**:129–36. [1597]
176. Nicol CF. A four year double blind study of tegretol in facial pain. *Headache* 1969;**9**:54–7. [1640]
177. Vilming ST, Lyberg T, Lataste X. Tizanidine in the management of trigeminal neuralgia. *Cephalalgia* 1986;**6**:181–2. [9143]
178. Lindström P, Lindblom U. The analgesic effect of tocainide in trigeminal neuralgia. *Pain* 1987;**28**:45–50. [3583]
179. Lechin F, van der Dijs B, Lechin ME, *et al*. Pimozide therapy for trigeminal neuralgia. *Arch Neurol* 1989;**46**:960–3. [1607]
180. Rull JA, Quibrera R, Gonzalez-Millan H, Lozano Castenada O. Symptomatic treatment of peripheral diabetic neuropathy with carbamazepine: double-blind crossover study. *Diabetologia* 1969;**5**:215–20. [1022]
181. Chadda VS, Mathur MS. Double blind study of the effects of diphenylhydantoin sodium on diabetic neuropathy. *J Assoc Physicians India* 1978;**26**:403–6. [1003]
182. Saudek CD, Werns S, Reidenberg MM. Phenytoin in the treatment of diabetic symmetrical polyneuropathy. *Clin Pharmacol Ther* 1977;**22**:196–9. [1646]
183. Rompel H, Bauermeister PW. Aetiology of migraine and prevention with carbamazepine (tegretol): results of a double-blind, cross-over study. *S Afr Med J* 1970;**44**:75–80. [9072]
184. Hering R, Kuritzky A. Sodium valproate in the prophylactic treatment of migraine: a double-blind study versus placebo. *Cephalalgia* 1992;**12**:81–4. [9070]
185. Stensrud P, Sjaastad O. Clonazepam (rivotril) in migraine prophylaxis. *Headache* 1979;**19**:333–4. [1650]
186. Greenbaum DS, Ferguson RK, Kater LA, Kuiper DH, Rosen LW. A controlled therapeutic study of the irritable-bowel syndrome. *N Engl J Med* 1973;**288**:13–16. [9071]
187. Leijon G, Boivie J. Central post-stroke pain – a controlled trial of amitriptyline and carbamazepine. *Pain* 1989;**36**:27–36. [1359]
188. Harkins S, Linford J, Cohen J, Kramer T, Cueva L. Administration of clonazepam in the treatment of TMD and associated myofascial pain: a double-blind pilot study. *J Craniomandib Disord* 1991;**5**:179–86. [1474]
189. Richards IM, Fraser SM, Hunter JA, Capell HA. Comparison of phenytoin and gold as second line drugs in rheumatoid arthritis. *Ann Rheum Dis* 1987;**46**:667–9. [1644]
190. Yajnik S, Singh GP, Singh G, Kumar M. Phenytoin as a coanalgesic in cancer pain. *J Pain Symptom Manag* 1992;**7**:209–13. [9069]
191. Gerson GR, Jones RB, Luscombe DK. Studies on the concomitant use of carbamazepine and clomipramine for the relief of post-herpetic neuralgia. *Postgrad Med J* 1977;**53**:104–9. [1371]
192. Martin C, Martin A, Rud C, Valli M. Comparative study of sodium valproate and ketoprofen in the treatment of postoperative pain. *Ann Fr Anesth Reanim* 1988;**7**:387–92. [1637]
193. Arieff AJ, Wetzel N. Tegretol in the treatment of neuralgias. *Dis Nerv Syst* 1967;**28**:820–3. [1304]
194. Fromm GH, Terrence CF, Chattha AS. Baclofen in the treatment of trigeminal neuralgia: double-blind study and long-term follow-up. *Ann Neurol* 1984;**15**:240–4. [930]
195. Goncikowska M. Treatment of Horton's headache with small doses of pilocarpine and carbamazepine. *Wiad Lek* 1984;**37**:1093–5. [1305]
196. Hatta V, Saxena A, Kaul HL. Phenytoin reduces suxamethonium-induced myalgia. *Anaesthesia* 1992;**47**:664–7. [1482]
197. Holmes B, Brogden RN, Heel RC, Speight TM, Avery GS. Flunarizine. A review of its pharmacodynamic and pharmacokinetic properties and therapeutic use. *Drugs* 1984;**27**:6–44. [1033]

198. Hopkins S. Clinical toleration and safety of azithromycin. *Am J Med* 1991;**91**:40S–45S. [1500]
199. Mathew NT, Ali S. Valproate in the treatment of persistent chronic daily headache. An open label study. *Headache* 1991;**31**:71–4. [1638]
200. Naidu MU, Ramesh Kumar T, Anuradha RT, Rao UR. Evaluation of phenytoin in rheumatoid arthritis: an open study. *Drugs Exp Clin Res* 1991;**17**:271–5. [1639]
201. Rasmussen P, Riishede J. Facial pain treated with carbamazepin (Tegretol). *Acta Neurol Scand* 1970;**46**:385–408. [1642]
202. Schaffler K, Wauschkuhn CH, Gierend M. Analgesic potency of a new anticonvulsant drug versus acetylsalicylic acid via laser somatosensory evoked potentials. Randomized placebo controlled double blind (5 way) crossover study. *Arzneimittelforschung* 1991;**41**:427–35. [1647]
203. Sharav Y, Benoliel R, Schnarch A, Greenberg L. Idiopathic trigeminal pain associated with gustatory stimuli. *Pain* 1991;**44**:171–4. [1648]
204. Shibasaki H, Tabira T, Inoue N, Goto I, Kuroiwa Y. Carbamazepine for painful crises in Fabry's disease. *J Neurol Sci* 1973;**18**:47–51. [1649]
205. Westerholm N. Treatment of facial pain with G 32 883 (Tegretol Geigy). *Scand J Dent Res* 1970;**78**:144–8. [1652]
206. Young RJ, Clarke BF. Pain relief in diabetic neuropathy: the effectiveness of imipramine and related drugs. *Diabetic Med* 1985;**2**:363–6. [12586]
207. Onghena P, van Houdenhove B. Antidepressant-induced analgesia in chronic non-malignant pain: a meta-analysis of 39 placebo-controlled studies. *Pain* 1992;**49**:205–19. [6430]
208. Gardner MJ, Altman DG. Statistics with confidence. London: British Medical Journal, 1989. [10715]
209. Sindrup SH, Gram LF, Broesen K, Eshoj O, Mogensen EF. The selective serotonin reuptake inhibitor paroxetine is effective in the treatment of diabetic neuropathy symptoms. *Pain* 1990;**42**:135–44. [3055]
210. McQuay HJ, Carroll D, Glynn CJ. Low dose amitriptyline in the treatment of chronic pain. *Anaesthesia* 1992;**47**:646–52. [4027]
211. Max MB, Lynch SA, Muir J, Shoaf SF, Smoller B, Dubner R. Effects of desipramine, amitriptyline, and fluoxetine on pain in diabetic neuropathy. *N Engl J Med* 1992;**326**:1250–6. [9242]
212. Spilker B. Guide to clinical trials. New York: Raven Press, 1991:699–700. [13205]
213. Davidoff G, Guarracini M, Roth E, Sliwa J, Yarkony G. Trazodone hydrochloride in the treatment of dysesthetic pain in traumatic myelopathy: a randomised, double-blind, placebo-controlled study. *Pain* 1987;**29**:151–61. [922]
214. Feinmann C. Psychogenic facial pain: presentation and treatment. *J Psychosom Res* 1983;**27**:403–10. [12386]
215. Feinmann C, Harris M. Psychogenic facial pain. Part 2: management and prognosis. *Br Dent J* 1984;**156**:205–8. [12396]
216. Feinmann C, Harris M, Cawley R. Psychogenic facial pain: presentation and treatment. *BMJ* 1984;**288**:436–8. [927]
217. Kvinesdal B, Molin J, Frøland A, Gram LF. Imipramin ved behandling af smerter ved diabetisk ekstremitetsneuropati. *Ugeskr Laeger* 1983;**145**:3018–19. [12426]
218. Kvinesdal B, Molin D, Frøland A, Gram LA. Imipramine therapy of painful diabetic neuropathy. *JAMA* 1984;**251**:1727–30. [947]
219. Mendel CM, Klein RF, Chappell DA, *et al*. A trial of amitriptyline and fluphenazine in the treatment of painful diabetic neuropathy. *JAMA* 1986;**255**:637–9. [10833]
220. Sharav Y, Singer E, Schmidt E, Dionne RA, Dubner R. The analgesic effect of amitriptyline on chronic facial pain. *Pain* 1991;**31**:199–209. [3305]
221. Sindrup SH, Gram LF, Skjold T, Frøland A, Beck-Nielsen H. Concentration-response relationship in imipramine treatment of diabetic neuropathy symptoms. *Clin Pharmacol Ther* 1990;**47**:509–15. [12536]
222. Sindrup SH, Bach FW, Gram LF. Plasma beta-endorphin is not affected by treatment with imipramine or paroxetine in patients with diabetic neuropathy symptoms. *Clin J Pain* 1992;**8**:145–8. [12506]
223. Stockstill JW, McCall WDJr, Gross AJ, Piniewski B. The effect of L tryptophan supplementation and dietary instruction on chronic myofascial pain. *J Am Dent Assoc* 1989;**118**:457–60. [12556]
224. Gomez-Perez FJ, Rull JA, Dies H, Rodriguez-Rivera JG, Gonzalez-Barranco J, Lozano-CastaOeda O. Nortriptyline and fluphenazine in the symptomatic treatment of diabetic neuropathy. A double blind cross over study. *Pain* 1985;**23**:395–400. [12406]
225. Max MB, Culnane M, Schafer SC, *et al*. Amitriptyline relieves diabetic neuropathy pain in patients with normal or depressed mood. *Neurology* 1987;**37**:589–96. [12456]
226. Max MB, Kishore-Kumar R, Schafer SC, *et al*. Efficacy of desipramine in painful diabetic neuropathy: a placebo-controlled trial. *Pain* 1991;**45**:69. [12466]
227. Sindrup SH, Ejlersen B, Frøland A, Sindrup EH, Brøsen K, Gram LF. Imipramine treatment in diabetic neuropathy: relief of subjective symptoms without changes in peripheral and autonomic nerve function. *Eur J Clin Pharmacol* 1989;**37**:151–3. [12596]

228. Sindrup SH, Gram LF, Skjold T, Grodum E, Brøsen K, Beck-Nielsen H. Clomipramine vs. desipramine vs. placebo in the treatment of diabetic neuropathy symptoms. A double-blind cross-over study. *Br J Clin Pharmacol* 1990;**30**:683–91. [12526]
229. Sindrup SH, Bjerre U, Dejgaard A, Brøsen K, Aaes-Jørgensen T, Gram LF. The selective serotonin reuptake inhibitor citalopram relieves the symptoms of diabetic neuropathy. *Clin Pharmacol Ther* 1992;**52**:547–52. [12516]
230. Sindrup SH, Tuxen C, Gram LF, *et al.* Lack of effect of mianserin on the symptoms of diabetic neuropathy. *Eur J Clin Pharmacol* 1992;**43**:251–5. [12546]
231. Kishore-Kumar R, Max MB, Schafer SC, *et al.* Desipramine relieves postherpetic neuralgia. *Clin Pharmacol Ther* 1990;**47**:305–12. [12416]
232. Max MB, Schafer SC, Culnane M, Smoller B, Dubner R, Gracely RH. Amitriptyline, but not lorazepam, relieves postherpetic neuralgia. *Neurology* 1988;**38**:1427–32. [12476]
233. Watson CP, Evans RJ, Reed K, Merskey H, Goldsmith L, Warsh J. Amitriptyline versus placebo in postherpetic neuralgia. *Neurology* 1982;**32**:671–3. [12576]
234. Lascelles RG. Atypical face pain and depression. *Br J Psychiatry* 1966;**112**:651–9. [12446]
235. Panerai AE, Monza G, Movilia P, Bianchi M, Francucci BM, Tiengo M. A randomized, within patient, cross over, placebo controlled trial on the efficacy and tolerability of the tricyclic antidepressants chlorimipramine and nortriptyline in central pain. *Acta Neurol Scand* 1990;**82**:34–8. [12486]
236. Langohr HD, Stöhr M, Petruch F. An open and double blind cross over study on the efficacy of clomipramine (Anafranil) in patients with painful mono- and polyneuropathies. *Eur Neurology* 1982;**21**:309–17. [12436]
237. Turkington RW. Depression masquerading as diabetic neuropathy. *JAMA* 1980;**243**:1147–50. [12606]
238. Watson CP, Chipman M, Reed K, Evans RJ, Birkett N. Amitriptyline versus maprotiline in postherpetic neuralgia: a randomized, double-blind, crossover trial. *Pain* 1992;**48**:29–36. [12566]
239. Burchiel KJ. Effects of electrical and mechanical stimulation on two foci of spontaneous activity which develop in primary afferent neurons after peripheral axotomy. *Pain* 1984;**18**:249–65. [1656]
240. Scadding JW. Development of ongoing activity, mechanosensitivity, and adrenaline sensitivity in severed peripheral nerve axons. *Exp Neurol* 1981;**72**:63–81. [1025]
241. Wall PD, Gutnick M. Properties of afferent nerve impulses originating from a neuroma. *Nature* 1974;**248**:740–3. [1657]
242. Wall PD, Devor M. Sensory afferent impulses originate from dorsal root ganglia as well as from the periphery in normal and nerve injured rats. *Pain* 1983;**17**:321–39. [1658]
243. Chabal C, Russell LC, Burchiel KJ. The effect of intravenous lidocaine, tocainide, and mexiletine on spontaneously active fibers originating in rat sciatic neuromas. *Pain* 1989;**38**:333–58. [4655]
244. Devor M, Wall PD, Catalan N. Systemic lidocaine silences ectopic neuroma and DRG discharge without blocking nerve conduction. *Pain* 1992;**48**:261–8. [8029]
245. Tanelian DL, MacIver MB. Analgesic concentrations of lidocaine suppress tonic A-delta and C fiber discharges produced by acute injury. *Anesthesiology* 1991;**74**:934–6. [4365]
246. Biella G, Sotgui ML. Central effects of systemic lidocaine mediated by glycine spinal receptors: an iontophoretic study in the rat spinal cord. *Brain Res* 1993;**603**:201–6. [1654]
247. Kastrup J, Angelo HR, Petersen P, Dejgaard A, Hilsted J. Treatment of chronic painful diabetic neuropathy with intravenous lidocaine infusion. *BMJ* 1986;**292**:173. [3588]
248. Petersen P, Kastrup J, Zeeberg I. Intravenous lidocaine i behandling af kroniske smerter [Intravenous lidocaine in the treatment of chronic pain]. *Ugeskr Laeger* 1986;**148**:2158–9. [1688]
249. Stracke H, Meyer U, Schumacher H, *et al.* Mexiletin in der Behandlung der schmerzhaften diabetischen Neuropathie [Mexiletine in treatment of painful diabetic neuropathy]. *Med Klin* 1994;**89**:124–31. [1690]
250. Petersen P, Kastrup J, Zeeberg I, Boysen G. Chronic pain treatment with intravenous lidocaine. *Neurol Res* 1986;**8**:189–90. [1691]
251. Galer BS, Harle J, Rowbotham MC. Response to intravenous lidocaine infusion predicts subsequent response to oral mexiletine: a prospective study. *J Pain Symptom Manag* 1996;**12**:161–7. [1684]
252. Wallace MS, Dyck JB, Rossi SS, Yaksh TL. Computer controlled lidocaine infusion for the evaluation of neuropathic pain after peripheral nerve injury. *Pain* 1996;**66**:69–77. [1696]
253. Dejgaard A, Petersen P, Kastrup J. Mexiletine for treatment of chronic painful diabetic neuropathy. *Lancet* 1988;**i**:9–11. [3745]
254. Kastrup J, Petersen P, Dejgaard A, Angelo HR, Hilsted J. Intravenous lidocaine infusion – a new treatment of chronic painful diabetic neuropathy? *Pain* 1987;**28**:69–75. [3589]
255. Stracke H, Meyer UE, Schumacher HE, Federlin K. Mexiletine in the treatment of diabetic neuropathy. *Diabetes Care* 1992;**15**:1550–5. [1689]

256. Rowbotham MC, Reisner-Keller LA, Fields HL. Both intravenous lidocaine and morphine reduce the pain of postherpetic neuralgia. *Neurology* 1991;**41**:1024–8. [3795]
257. ChiouTan FY, Tuel SM, Johnson JC, Priebe MM, Hirsh DD, Strayer JR. Effect of mexiletine on spinal cord injury dysesthetic pain. *Am J Phys Med Rehabil* 1996;**75**:84–7. [1692]
258. Sörensen J, Bengtsson A, Bäckman E, Henriksson KG, Bengtsson M. Pain analysis in patients with fibromyalgia. *Scand J Rheumatol* 1995;**24**:360–5. [1697]
259. Marbach JJ, Wallenstein SL. Analgesic, mood, and hemodynamic effects of intranasal cocaine and lidocaine in chronic facial pain of deafferentation and myofascial origin. *J Pain Symptom Manag* 1988;**3**:73–9. [815]
260. Bruera E, Ripamonti C, Brenneis C, Macmillan K, Hanson J. A randomized double-blind crossover trial of intravenous lidocaine in the treatment of neuropathic cancer pain. *J Pain Sympt Manag* 1992;**7**:138–40. [1682]
261. Ellemann K, Sjøgren P, Banning A-M, Jensen TS, Smith T, Geertsen P. Trial of intravenous lidocaine on painful neuropathy in cancer patients. *Clin J Pain* 1989;**5**:291–4. [1687]
262. Sjøgren P, Banning AM, Hebsgaard K, Petersen P, Gefke K. Intravenøs lidokain i behandlingen af kroniske smerter forårsaget af knoglemetastaser [Intravenous lidocaine in the treatment of chronic pain caused by bone metastases]. *Ugeskr Laeger* 1989;**151**:2144–6. [1686]
263. Bell R, Montoya D, Shuaib A, Lee MA. A comparative trial of three agents in the treatment of acute migraine headache. *Ann Emerg Med* 1990;**19**:1079–82. [1683]
264. Maizels M, Scott B, Cohen W, Chen W. Intranasal lidocaine for treatment of migraine: a randomized, double-blind, controlled trial. *JAMA* 1996;**276**:319–21. [1583]
265. Reutens DC, Fatovich DM, Stewart Wynne EG, Prentice DA. Is intravenous lidocaine clinically effective in acute migraine? *Cephalalgia* 1991;**11**:245–7. [1693]
266. Chabal C, Jacobson L, Mariano A, Chaney E, Britell CW. The use of oral mexiletine for the treatment of pain after peripheral nerve injury. *Anesthesiology* 1992;**76**:513–17. [6612]
267. Dunlop R, Davies RJ, Hockley J, Turner P. Analgesic effects of oral flecainide. *Lancet* 1988;**i**:420–1. [727]
268. Tfelt-Hansen P. Sumatriptan for the treatment of migraine attacks—a review of controlled clinical studies. *Cephalalgia* 1993;**13**:238–44. [1589]
269. 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. [12028]
270. Rational use of NSAIDs for musculoskeletal disorders. *Drug Ther Bull* 1994;**32**(12):91–5. [12029]
271. Mondino A, Zanolò G, Giachetti C, Testaguzza F, Engels B, Wagener HH. Humankinetische Untersuchungen mit Ibuprofen. Vergleichende Studie zur Bestimmung der Serumkonzentrationen und der Metaboliten nach lokaler und oraler Gabe. *Med Welt* 1983;**34**(38):57–9. [12032]
272. Berner G, Engels B, Vögtle-Junkert U. Percutaneous ibuprofen therapy with Trauma-Dolgit gel: bioequivalence studies. *Drugs Exp Clin Res* 1989;**15**:559–64. [12033]
273. Treffel P, Gabard B. Feasibility of measuring the bioavailability of topical ibuprofen in commercial formulations using drug content in epidermis and a methyl nicotinate skin inflammation assay. *Skin Pharmacol* 1993;**6**:268–75. [12034]
274. Wynne HA, Campbell M. Pharmacoeconomics of nonsteroidal anti-inflammatory drugs (NSAIDs). *PharmacoEcon* 1993;**3**:107–23. [12036]
275. McKell D, Stewart A. A cost-minimisation analysis comparing topical versus systemic NSAIDs in the treatment of mild osteoarthritis of the superficial joints. *Br J Med Econ* 1994;**7**:137–46. [12035]
276. Evans JMM, McMahon AD, McGilchrist MM, *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. [12046]
277. Fordyce WE. Behavioral methods for chronic pain and illness. St Louis, Missouri: CV Mosby, 1976.
278. Rudy TE, Kerns RD, Turk DC. Chronic pain and depression: toward a cognitive-behavioral mediation model. *Pain* 1988;**35**:129–140.
279. Gamsa A. The role of psychological factors in chronic pain. I. A half century of study. *Pain* 1994;**57**:5–15.
280. Gil KM, Ross SL, Keefe FJ. Behavioral treatment of chronic pain: four pain management protocols. In: France RD, Krishnan KDD, editors. Chronic pain. Washington: American Psychiatric Press, 1988:376–413.
281. Pearce S. A review of cognitive-behavioural methods for the treatment of chronic pain. *J Psychosom Res* 1983;**27**:431–40.
282. Eccleston C. The attentional control of pain: methodological and theoretical concerns. *Pain* 1995;**63**:3–10.
283. Dolce JJ. Self-efficacy and disability beliefs in behavioral treatment of pain. *Behav Res Ther* 1987;**25**:289–99.

284. Rosenstiel AK, Keefe FJ. The use of coping strategies in chronic low back pain patients: relationship to patient characteristics and current adjustment. *Pain* 1983;**17**:33–44.
285. Jensen MP, Turner JA, Romano JM. Self-efficacy and outcome expectancies: relationship to chronic pain, coping strategies and adjustment. *Pain* 1991;**44**:263–9.
286. Beck AT, Rush AJ, Shaw BF, Emery G. Cognitive therapy of depression. New York: Guilford Press, 1979.
287. American Psychiatric Association. Diagnostic and statistical manual of mental disorders. 4th edition: DSMIV. Washington: American Psychiatric Association, 1994.
288. Turk DC, Salovey P. "Chronic pain as a variant of depressive disease": a critical reappraisal. *J Nerv Ment Dis* 1984;**172**:398–404.
289. Turk DC, Meichenbaum D, Genest M. Pain and behavioral medicine: a cognitive-behavioral perspective. New York: The Guilford Press, 1983.
290. Linton SJ. Behavioral remediation of chronic pain: a status report. *Pain* 1986;**24**:125–41.
291. Lewinsohn PM. The behavioral study and treatment of depression. In: Hersen M, Eisler RM, Miller PM, editors. Progress in behavior modification. New York: Academic Press, 1975: 1.
292. Linton SJ. A critical review of behavioural treatments for chronic benign pain other than headache. *Br J Clin Psychol* 1982;**21**:321–37.
293. Block AR. Multidisciplinary treatment of chronic low back pain: a review. *Rehab Psychology* 1982;**27**:51–63.
294. Keefe FJ. Behavioral assessment and treatment of chronic pain: current status and future directions. *J Consult Clin Psychol* 1982;**50**:886–91.
295. Aronoff GM, Evans WO, Enders PL. A review of follow-up studies of multidisciplinary pain units. *Pain* 1983;**16**:1–11.
296. Trifiletti RJ. The psychological effectiveness of pain management procedures in the context of behavioral medicine and medical psychology. *Genet Psychol Monographs* 1984;**109**:251–78.
297. Blackwell B, Galbraith JR, Dahl DS. Chronic pain management. *Hosp Commun Psychiatry* 1984;**35**:999–1008.
298. Turk DC, Flor H. Etiological theories and treatments for chronic back pain. II. Psychological models and interventions. *Pain* 1984;**19**:209–33.
299. Follick MJ, Ahern DK, Attanasio V, Riley JF. Chronic pain programs: current aims, strategies, and needs. *Ann Behav Med* 1985;**7**:17–20.
300. Merskey H. Chronic pain syndromes and their treatment. In: Kennard C, editor. Recent advances in clinical neurology, No. 5. Edinburgh: Churchill Livingstone, 1988:87–107.
301. Benjamin S. Psychological treatment of chronic pain: a selective review. *J Psychosom Res* 1989;**33**:121–31.
302. Linssen ACG, Spinhoven P. Multimodal treatment programmes for chronic pain: a quantitative analysis of existing research data. *J Psychosom Res* 1992;**36**:275–86.
303. Keefe FJ, Dunsmore J, Burnett R. Behavioral and cognitive-behavioral approaches to chronic pain: recent advances and future directions. *J Consult Clin Psychology* 1992;**60**:528–36.
304. Nicholas MK. Chronic pain. In: Wilson PH, editor. Principles and practice of relapse prevention. New York: Guildford Press, 1992:255–89.
305. Malone MD, Strube MJ. Meta-analysis of non-medical treatments for chronic pain. *Pain* 1988;**34**:231–44.
306. Flor H, Fydrich T, Turk DC. Efficacy of multidisciplinary pain treatment centers: a meta-analytic review. *Pain* 1992;**49**:221–30.
307. Fishbain DA, Large RG, Lieber SJ. Evaluation of multidisciplinary pain clinics [abstract]. In: 8th World Congress on Pain Abstract Book. Seattle: IASP Press, 1996: 101.
308. Williams AC de C. Cognitive-behavioural management of chronic pain: models and outcomes. [PhD thesis: unpublished]. London: University of London, 1996.
309. International Association for the Study of Pain. Desirable characteristics for pain treatment facilities. Seattle: IASP, 1990.
310. Shapiro DA, Shapiro D. Meta-analysis of comparative outcome studies: a replication and refinement. *Psychol Bull* 1982;**92**:581–604.
311. Roth AD, Fonagy P. What works for whom? A critical review of psychotherapy research. New York: Guildford Press, 1996.
312. Turk DC. Efficacy of multidisciplinary pain centers in the treatment of chronic pain. In: Cohen MJM, Campbell JN, editors. Pain centers at the crossroads: progress in pain research and management. Seattle: IASP Press, 1996: 7.
313. Turk DC, Meichenbaum DH, Berman WH. Application of biofeedback for the regulation of pain: a critical review. *Psychol Bull* 1979;**86**:1322–38.
314. Turner JA, Chapman CR. Psychological interventions for chronic pain: a critical review. I. Relaxation training and biofeedback. *Pain* 1982;**12**:1–21.

315. Lorig K, Fries JF. The arthritis helpbook. Reading MA: Addison-Wesley, 1980.
316. Borkovec TD, Nau SD. Credibility of analogue therapy rationales. *J Behav Ther Exp Psychiatry* 1972;**3**:257–60.
317. Meenan RF. AIMS approach to health status measurement: conceptual background and measurement properties. *J Rheumatol* 1982;**9**:785–8.
318. Keefe FJ, Block AR. Development of an observation method for assessing pain behavior in chronic low back pain patients. *Behav Ther* 1982;**13**:363–75.
319. Nicholas MK. An evaluation of treatments for chronic low back pain [Doctoral dissertation: unpublished]. Sydney, Australia; University of Sydney, 1988.
320. Harding VR, Williams Acde C, Richardson PH, *et al*. The development of a battery of measures for assessing physical functioning of chronic pain patients. *Pain* 1994;**58**:367–75.
321. Williams AC de C. NNTs used in decision-making in chronic pain management. *Bandolier* 1995;**2**:3.
322. Pither CE, Ralphs JA. Limiting the drugs list: behavioural treatment [letter]. *BMJ* 1993;**306**:1687–8.
323. Altman DG. Better reporting of randomised controlled trials: the CONSORT statement. *BMJ* 1996;**313**:570–1.
324. Clarke MJ, Stewart LA. Obtaining data from randomised controlled trials: how much do we need for reliable and informative meta-analyses? *BMJ* 1994;**309**:1007–10.
325. Melzack R. The McGill Pain Questionnaire: major properties and scoring methods. *Pain* 1975;**1**:277–99.
326. Kerns RD, Turk DC, Rudy TE. The West Haven-Yale Multidimensional Pain Inventory (WHYMPI). *Pain* 1985;**23**:345–56.
327. Bergner M, Bobbitt RA, Carter WB, Gilson BS. The Sickness Impact Profile: development and final revision of a health status measure. *Med Care* 1981;**19**:787–805.
328. Radloff L. The CES-D scale: a self-report depression scale for research in the general population. *Appl Psychol Measure* 1977;**1**:385–401.
329. Spielberger CD. Manual for the State-Trait Anxiety Inventory. Palo Alto, CA: Consulting Psychologists Press, 1983.
330. Gottlieb BS. Development of the Pain Beliefs Questionnaire: a preliminary report. In: Proceedings of Association for the Advancement of Behaviour Therapy; Philadelphia PA, 1984.
331. Vlaeyen JWS, Geurts SM, van Eek H, Kole-Snijders AMJ, Schuerman JA, Groenman NH. What do chronic pain patients think of their pain: development of a pain cognition questionnaire. *Br J Clin Psychol* 1990;**29**:383–94.
332. Andrews G. Talk that works: the rise of cognitive behaviour therapy. *BMJ* 1996;**313**:1501–2.
333. McQuay HJ, Machin L, Moore RA. Chronic non-malignant pain: a population prevalence study. *Practitioner* 1985;**229**:1109–11. [769]
334. Barsky AJ, Borus JF. Somatization and medicalization in the era of managed care. *JAMA* 1995;**274**:1931–4. [1694]
335. Weir R, Browne GB, Tunks E, Gafni A, Roberts J. A profile of users of speciality pain clinic services: predictors of use and cost estimates. *J Clin Epidemiol* 1992;**45**:1399–1415. [10123]
336. Nash TP, Williams JD, Machin D. TENS: does the type of stimulus really matter? *Pain Clinic* 1990;**3**:161–8.
337. Farago F. Trigeminal neuralgia: its treatment with two new carbamazepine analogues. *Eur Neurol* 1987;**26**:73–83.

Systematic 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.

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

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

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.

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

Baron JC, Boussier 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.

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.

- Bernstein J, 1983. Overview of efficacy of fenbufen in rheumatoid arthritis and osteoarthritis. *Am J Med*;75:70–4.
- Bhatt Sanders D, 1985. Acupuncture for rheumatoid arthritis: an analysis of the literature. *Semin Arthritis Rheum*;14:225–31.
- 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.
- Blanchard EB, Andrasik F, 1982. Psychological assessment and treatment of headache: recent developments and emerging issues. *J Consult Clin Psychol*;50:859–79.
- Bogaards MC, ter Kuile MM, 1994. Treatment of recurrent tension headache: a meta-analytic review. *Clin J Pain*;10:174–90.
- Bosi Ferraz M, Atra E, 1989. Metaanálise: nova opção em pesquisa clínica. Princípios 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.
- 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.
- 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.
- Crooks RJ, Jones Da, Fiddian P, 1991. Zoster-associated chronic pain: an overview of clinical trials with acyclovir. *Scand J Infect*;78:62–8.
- 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 SL, 1993. Nabumetone: a “nonacidic” nonsteroidal anti-inflammatory drug. *Ann Pharmacother*;27:456–63.
- Di Fabio RP, 1992. Efficacy of manual therapy. *Phys Ther*;72:853–64.
- 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.
- Ernst E, Fialka V, 1993. Low-Dose-Lasertherapie: eine kritische Prüfung der klinischen Wirksamkeit [Low-dose laser therapy: critical analysis of clinical effect]. *Schweiz Med Wochenschr*;123:949–54.
- 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, Nielsen GL, Schmidt EB, 1994. Fiskeolie og reumatoid artrit [Fish oils and rheumatoid arthritis]. *Ugeskr-Laeger*;156:3490–5.
- Felson DT, Anderson JJ, Boers M, *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, 1990. The comparative efficacy and toxicity of second-line drugs in rheumatoid arthritis. *Arthritis Rheum*;33:1449–61.
- 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.
- 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.
- Fernandez E, Turk DC, 1989. The utility of cognitive coping strategies for altering pain perception: a meta analysis. *Pain*;38:123–35.
- Ferraz MB, Tugwell P, Goldsmith CH, Atra E, 1990. Meta-analysis of sulfasalazine in ankylosing spondylitis. *J Rheumatol*;17:1482–6.
- Flor H, Fydrich T, Turk DC, 1992. Efficacy of multidisciplinary 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.
- 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*;96:33–6.
- Glazer S, Portenoy RK, 1991. Systemic local anesthetics in pain control. *J Pain Symptom Manag*;6:30–9.
- 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, 1993. Meta-analysis of NSAIDs: contribution of drugs, doses, trial designs, and meta-analytic techniques. *Scand J Rheumatol*;22:255–60.
- 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;10:356].
- 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, 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.
- Hill AG, 1967. Review of flufenamic acid in rheumatoid arthritis. *Ann Phys Med*;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, Athens OH, Penzien DB, Jackson MS, Rokicki La, Cordingley GE, 1992. Flunarizine vs propranolol: a meta-analysis of clinical trials. *Headache*;32:256.
- 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, 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.
- 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, 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, McQuay HJ, 1996. Meta-analyses to evaluate analgesic interventions: a systematic qualitative review of their methodology. *J Clin Epidemiol*;49:235–43.
- Kaiko RF, Grandy RP, Oshlack B, *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, Tramer M, Carroll D, McQuay H, Moore RA, 1996. Pain relief from intra-articular morphine after knee surgery: a qualitative systematic review. *Pain*. In press.
- Kellihan MJ, Mangino PD, 1992. Pamidronate. *Ann Pharmacother*;26:1262–9.
- 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*;303:1298–1303.
- 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, 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, 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.
- 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.
- 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.
- 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.
- Magni G, 1991. The use of antidepressants in the treatment of chronic pain. A review of the current evidence. *Drugs*;42:730–48.
- Malone MD, Strube MJ, 1988. Meta-analysis of non-medical treatments for chronic pain. *Pain*;34:231–44. [erratum: *Pain* 1989;37:128].
- 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 HJ, Carroll D, Glynn CJ, 1993. Dose-response for analgesic effect of amitriptyline in chronic pain. *Anaesthesia*;48:281–5.
- McQuay H, Moore RA, 1996. Epidural steroids [letter]. *Anaesth Intensive Care*;24:284–6.
- McQuay H, Nye BA, Carroll D, Wiffen PJ, Tramèr M, Moore RA, 1996. A systematic review of antidepressants in neuropathic pain. *Pain* 1996;68:217–27.
- 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.
- 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 suppl: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.
- 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, DiFabio RP, 1985. Efficacy of spinal manipulation/mobilization therapy. A meta analysis. *Spine*;10:833–7.
- 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, 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, Holroyd KA, 1987. The behavioral treatment of migraine headache: a meta-analytic review of the literature. *Dissertation Abstracts Int*;47: No DA8629941.
- 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.
- 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.
- 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 hydroxyethylrutosides in the treatment of chronic venous insufficiency. *Vasa*;23:244–50.
- Pradalier A, Vincent D, 1992. Migraine et anti-inflammatoires non-steroidiens. *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.
- 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, Wyderski RJ, 1990. Conservative management of intermittent claudication. *Ann Intern Med*;113:135–46.
- 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*;20:2095–2103.
- 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*;43:610–5.

- Schiassi M, Bianchini C, Restelli A, Zoni G, 1989. Tolerability profile of the anti-inflammatory compound imidazole salicylate: a meta-analysis of safety data in 1408 patients. *Int J Tiss React*;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.
- Shekelle PG, Adams AH, Chassin MR, Hurwitz EL, Brook RH, 1992. Spinal manipulation for low-back pain. *Ann Intern Med*;117:590–8.
- Simmons J, Stavinoha W, Knodel L, 1984. Update and review of chemonucleolysis. *Clin Orthop*;Mar:51–60.
- 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.
- 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.
- 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.
- Trabant H, Widdra W, De Looze S, 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, 1996. Keep taking the tablets! A systematic review of direct comparisons of non-steroidal anti-inflammatory drugs given by different routes for acute pain. Submitted.
- 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, 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 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.
- van der Heijden CJ, 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.
- 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.
- 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.
- 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.

Psychological reviews and RCTs

- Altmaier EM, Lehmann TR, Russell DW, Weinstein JN, Kao CF, 1992. The effectiveness of psychological interventions for the rehabilitation of low back pain: a randomized controlled trial evaluation. *Pain*;49:329–35.
- Appelbaum KA, Blanchard EB, Hickling EJ, Alfonso M, 1988. Cognitive behavioral treatment of a veteran population with moderate to severe rheumatoid arthritis. *Behavior Therapy*;19:489–502.
- Bradley LA, Young LD, Anderson KO, *et al*, 1987. Effects of psychological therapy on pain behavior of rheumatoid arthritis patients. *Arthritis Rheum*;30:1105–14.
- Cott A, McCully J, Goldberg WM, Tanser PH, Parkinson W, 1992. Interdisciplinary treatment of morbidity in benign chest pain. *Angiology*;195–202.
- Flor H, Birbaumer N, 1993. Comparison of the efficacy of electromyographic biofeedback, cognitive-behavioral therapy, and conservative medical interventions in the treatment of chronic musculoskeletal pain. *J Consult Clin Psychology*;61:653–8.

- Keefe FJ, Caldwell DS, Williams DA, *et al*, 1990. Pain coping skills training in the management of osteoarthritic knee pain: a comparative study. *Behav Ther*;21:49–63.
- Keefe FJ, Caldwell DS, Baucom D, *et al*, 1996. Spouse-assisted coping skills training in the management of osteoarthritic knee pain. *Arthritis Care Res*;9:279–91.
- Keefe FJ, Caldwell DS, Williams DA, *et al*, 1990. Pain coping skills training in the management of osteoarthritic knee pain – II: follow-up results. *Behav Ther*;21:453–7.
- Kerns RD, Turk DC, Holzman AD, Rudy TE, 1986. Comparison of cognitive-behavioral and behavioral approaches to the outpatient treatment of chronic pain. *Clin J Pain*;1:195–203.
- Kraaimaat FW, Brons MR, Geenen R, Bijlsma JWJ, 1995. The effect of cognitive behavior therapy in patients with rheumatoid arthritis. *Behav Res Ther*;33:487–95.
- Linton SJ, Gotestam KG, 1984. A controlled study of the effects of applied relaxation and applied relaxation plus operant procedures in the regulation of chronic pain. *Br J Clin Psychol*;23:291–9.
- Linton SJ, Melin L, Stjernlof K, 1985. The effects of applied relaxation and operant activity training on chronic pain. *Behav Psychother*;13:87–100.
- Moore JE, Chaney EF, 1985. Outpatient group treatment of chronic pain: effects of spouse involvement. *J Consult Clin Psychol*;53:326–44.
- Newton-John TRO, Spence SH, Schotte D, 1995. Cognitive-behavioural therapy versus EMG biofeedback in the treatment of chronic low back pain. *Behav Res Ther*;33:691–7.
- Nicholas MK, Wilson PH, Goyen J, 1991. Operant-behavioural and cognitive-behavioural treatment for chronic low back pain. *Behav Res Ther*;29:225–38.
- Nicholas MK, Wilson PH, Goyen J, 1992. Comparison of cognitive-behavioral group treatment and an alternative non-psychological treatment for chronic low back pain. *Pain*;48:339–47.
- O'Leary A, Shoor S, Lorig K, Holman HR, 1988. A cognitive-behavioral treatment for rheumatoid arthritis. *Health Psychol*;7:527–44.
- Parker JC, Frank RG, Beck NC, *et al*, 1988. Pain management in rheumatoid arthritis patients. A cognitive-behavioral approach. *Arthritis Rheum*;31:593–601.
- Peters AAW, van Dorst E, Jellis B, van Zuuren E, Hermans J, Trimbos JB, 1991. A randomized clinical trial to compare two different approaches in women with chronic pelvic pain. *Obstetr Gynecol*;77:740–4.
- Peters JL, Large RG, 1990. A randomised control trial evaluating in- and outpatient pain management programmes. *Pain*;41:283–93.
- Peters JL, Large RG, Elkind G, 1992. Follow-up results from a randomised controlled trial evaluating in- and outpatient pain management programmes. *Pain*;50:41–50.
- Pilowsky I, Spence N, Rounsefell B, Forsten C, Soda J, 1995. Out-patient cognitive-behavioural therapy with amitriptyline for chronic non-malignant pain: a comparative study with 6-month follow-up. *Pain*;60:49–54.
- Puder RS, 1988. Age analysis of cognitive-behavioral group therapy for chronic pain outpatients. *Psychol Aging*;3:204–7.
- Radojevic V, Nicassio PM, Weisman MH, 1992. Behavioral intervention with and without family support for rheumatoid arthritis. *Behav Ther*;23:13–30.
- Spence SH, 1989. Cognitive-behavior therapy in the management of chronic, occupational pain of the upper limbs. *Behav Res Ther*;27:435–66.
- Spence SH, 1991. Cognitive-behavior therapy in the management of chronic, occupational pain of the upper limbs: a 2 year follow-up. *Behav Res Ther*;29:503–9.
- Spence SH, Sharpe L, Newton-John T, Champion D, 1995. Effect of EMG biofeedback compared to applied relaxation training with chronic upper extremity cumulative trauma disorders. *Pain*;63:199–203.
- Strauss GD, Spiegel JS, Daniels M, *et al*, 1986. Group therapies for rheumatoid arthritis: a controlled study of two approaches. *Arthritis Rheum*;29:1203–9.
- Turner JA, 1982. Comparison of group progressive-relaxation training and cognitive-behavioral group therapy for chronic low back pain. *J Consult Clin Psychol*;50:757–65.
- Turner JA, Clancy S, 1988. Comparison of operant behavioral and cognitive-behavioral group treatment for chronic low back pain. *J Consult Clin Psychol*;56:261–6.
- Turner JA, Clancy S, McQuade KJ, Cardenas DD, 1990. Effectiveness of behavioral therapy for chronic low back pain: a component analysis. *J Consult Clin Psychol*;58:573–9.
- Turner JA, Jensen MP, 1993. Efficacy of cognitive therapy for chronic low back pain. *Pain*;52:169–77.
- Vlaeyen JWS, Haazen IWCJ, Kole-Snijders AMJ, van Eek H, Schuerman JA, 1995. Behavioral rehabilitation of chronic low back pain: comparison of an operant treatment, an operant-cognitive treatment and an operant-respondent treatment. *Br J Clin Psychol*;34:95–118.
- Vlaeyen JWS, Teeken-Gruben NJG, Goossens MEJB, *et al*, 1996. Cognitive-educational treatment of fibromyalgia: a randomized clinical trial. I. Clinical effects. *J Rheumatol*;23:1237–45.
- Williams AC de C, Richardson PH, Nicholas MK, *et al*, 1996. Inpatient vs. outpatient pain management: results of a randomised controlled trial. *Pain*;66:13–22.

Acute Sector Panel

Chair: Professor John Farndon, University of Bristol †

Professor Senga Bond, University of Newcastle-upon-Tyne †	Professor Richard Ellis, St James's University Hospital, Leeds †	Dr Chris McCall, General Practitioner, Dorset †	Professor Jon Nicoll, University of Sheffield †	Dr William Tarnow-Mordi, University of Dundee
Professor Ian Cameron, SE Thames RHA	Dr David Field, Leicester Royal Infirmary NHS Trust †	Professor Alan McGregor, St Thomas's Hospital, London	Professor John Norman, Southampton University	Professor Kenneth Taylor, Hammersmith Hospital, London †
Ms Lynne Clemence, Mid-Kent Health Care Trust †	Mr Ian Hammond, Hillingdon HA †	Mrs Wilma MacPherson, St Thomas's & Guy's Hospitals, London	Professor Gordon Stirrat, St Michael's Hospital, Bristol	
Professor Cam Donaldson, University of Aberdeen †	Professor Adrian Harris, Churchill Hospital, Oxford		Professor Michael Sheppard, Queen Elizabeth Hospital, Birmingham †	

Diagnostics and Imaging Panel

Chair: Professor Mike Smith, University of Leeds †

Professor Michael Maisey, Guy's & St Thomas's Hospitals, London*	Professor MA Ferguson-Smith, University of Cambridge †	Professor Donald Jeffries, St Bartholomew's Hospital, London †	Professor Colin Roberts, University of Wales College of Medicine †	Mr Stephen Thornton, Cambridge & Huntingdon Health Commission
Professor Andrew Adam, UMDS, London †	Dr Mansel Hickey, University of Manchester	Dr Andrew Moore, Editor, Bandolier †	Miss Annette Sergeant, Chase Farm Hospital, Enfield	Dr Gillian Vivian, Royal Cornwall Hospitals Trust †
Dr Pat Cooke, RDRD, Trent RHA	Professor Sean Hilton, St George's Hospital Medical School, London	Professor Chris Price, London Hospital Medical School †	Professor John Stuart, University of Birmingham	Dr Jo Walsworth-Bell, South Staffordshire Health Authority †
Ms Julia Davison, St Bartholomew's Hospital, London †	Mr John Hutton, MEDTAP Europe Inc., London †	Dr Ian Reynolds, Nottingham HA	Dr Ala Szczepura, University of Warwick †	Dr Greg Warner, General Practitioner, Hampshire †

Methodology Panel

Chair: Professor Anthony Culyer, University of York †

Mr Doug Altman, Institute of Health Sciences, Oxford †	Dr Rory Collins, University of Oxford	Dr Stephen Harrison, University of Leeds	Professor Ian Russell, University of York †	Dr David Spiegelhalter, Institute of Public Health, Cambridge †
Professor Michael Baum, Royal Marsden Hospital	Professor George Davey-Smith, University of Bristol	Mr Philip Hewitson, Leeds FHSA	Professor David Sackett, Centre for Evidence Based Medicine, Oxford †	Professor Charles Warlow, Western General Hospital, Edinburgh †
Professor Nick Black, London School of Hygiene & Tropical Medicine †	Professor Ray Fitzpatrick, University of Oxford †	Professor Richard Lilford, Regional Director, R&D, West Midlands †	Dr Maurice Slevin, St Bartholomew's Hospital, London	
Professor Martin Buxton, Brunel University †	Professor Stephen Frankel, University of Bristol	Mr Nick Mays, Kings Fund Institute, London †		

Pharmaceutical Panel

Chair: Professor Tom Walley, University of Liverpool †

Professor Michael Rawlins, University of Newcastle-upon-Tyne*	Ms Christine Clarke, Hope Hospital, Salford †	Dr Desmond Fitzgerald, Mere, Bucklow Hill, Cheshire †	Professor Trevor Jones, ABPI, London †	Dr Ross Taylor, University of Aberdeen †
Dr Colin Bradley, University of Birmingham	Mrs Julie Dent, Ealing, Hammersmith and Hounslow HA, London †	Dr Alistair Gray, Wolfson College, Oxford †	Dr Andrew Mortimore, Southampton & SW Hants Health Authority †	Dr Tim van Zwanenberg, Northern RHA
Professor Alasdair Breckenridge, RDRD, Northwest RHA	Mr Barrie Dowdeswell, Royal Victoria Infirmary, Newcastle-upon-Tyne	Professor Keith Gull, University of Manchester	Dr John Posnett, University of York	Dr Kent Woods, RDRD, Trent RO, Sheffield †
		Dr Keith Jones, Medicines Control Agency	Dr Frances Rotblat, Medicines Control Agency †	

Population Screening Panel

Chair: Professor Sir John Grimley Evans, Radcliffe Infirmary, Oxford †

Dr Sheila Adam, Department of Health*	Professor George Freeman, Charing Cross & Westminster Medical School, London	Dr Ann Ludbrook, University of Aberdeen †	Professor Catherine Peckham, Institute of Child Health, London †	Professor Nick Wald, University of London †
Dr Anne Dixon Brown, NHS Executive, Anglia & Oxford†	Dr Mike Gill, Brent & Harrow Health Authority †	Professor Alexander Markham, St James's University Hospital, Leeds †	Dr Connie Smith, Parkside NHS Trust, London †	Professor Ciaran Woodman, Centre for Cancer Epidemiology, Manchester †
Professor Dian Donnai, St Mary's Hospital, Manchester †	Dr JA Muir Gray, RDRD, Anglia & Oxford RO †	Professor Theresa Marteau, UMDS, London †	Dr Sarah Stewart-Brown, University of Oxford †	

Primary and Community Care Panel

Chair: Professor Angela Coulter, Kings Fund Centre for Health Services Development, London †

Professor Martin Roland, University of Manchester*	Professor Andrew Haines, RDRD, North Thames RHA	Professor Roger Jones, UMDS, London †	Dr Fiona Moss, North Thames British Postgraduate Medical Federation†	Dr Mary Renfrew, University of Oxford
Dr Simon Allison, University of Nottingham	Dr Nicholas Hicks, Oxfordshire Health Authority †	Mr Lionel Joyce, Chief Executive, Newcastle City Health NHS Trust †	Professor Dianne Newham, Kings College, London	Dr John Tripp, Royal Devon & Exeter Healthcare NHS Trust †
Mr Kevin Barton, Bromley Health Authority †	Professor Richard Hobbs, University of Birmingham †	Professor Martin Knapp, London School of Economics & Political Science †	Professor Gillian Parker, University of Leicester †	
Professor John Bond, University of Newcastle-upon-Tyne †	Professor Allen Hutchinson, University of Hull †	Professor Karen Luker, University of Liverpool	Dr Robert Peveler, University of Southampton †	* Previous Chair † Current members
Professor Shah Ebrahim, Royal Free Hospital, London	Mr Edward Jones, Rochdale FHSA			

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