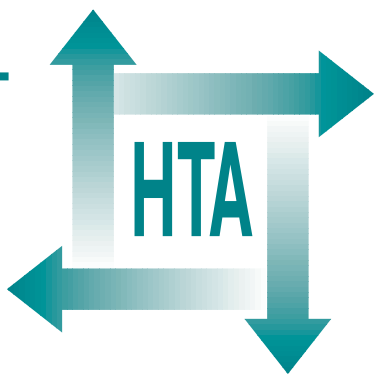
Health Technology Assessment 2001; Vol. 5: No. 16

Methodology

How to develop cost-conscious guidelines

M Eccles
J Mason

Health Technology Assessment
NHS R&D HTA Programme



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.

How to develop cost-conscious guidelines

M Eccles[1*]
J Mason[2]

[1] Centre for Health Services Research, University of Newcastle, Newcastle upon Tyne, UK
[2] Centre for Health Economics, University of York, UK

* Corresponding author

Competing interests: James Mason has received academic grants, fees and expenses for research and consultancy work from pharmaceutical companies that manufacture treatments discussed in this report.

Published May 2001

This report should be referenced as follows:

Eccles M, Mason J. How to develop cost-conscious guidelines. *Health Technol Assess* 2001;**5**(16).

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

NHS R&D HTA Programme

The NHS R&D Health Technology Assessment (HTA) Programme was set up in 1993 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 provide care in the NHS.

Initially, six HTA panels (pharmaceuticals, acute sector, primary and community care, diagnostics and imaging, population screening, methodology) helped to set the research priorities for the HTA Programme. However, during the past few years there have been a number of changes in and around NHS R&D, such as the establishment of the National Institute for Clinical Excellence (NICE) and the creation of three new research programmes: Service Delivery and Organisation (SDO); New and Emerging Applications of Technology (NEAT); and the Methodology Programme.

Although the National Coordinating Centre for Health Technology Assessment (NCCHTA) commissions research on behalf of the Methodology Programme, it is the Methodology Group that now considers and advises the Methodology Programme Director on the best research projects to pursue.

The research reported in this monograph was funded as project number 94/08/28.

The views expressed in this publication are those of the authors and not necessarily those of the Methodology Programme, HTA Programme or the Department of Health. The editors wish to emphasise that funding and publication of this research by the NHS should not be taken as implicit support for any recommendations made by the authors.

Criteria for inclusion in the HTA monograph series
Reports are published in the HTA monograph series if (1) they have resulted from work commissioned for the HTA Programme, and (2) they are of a sufficiently high scientific quality as assessed by the referees and editors.

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

The editors and publisher have tried to ensure the accuracy of this report but do not accept liability for damages or losses arising from material published in this report. They would like to thank the referees for their constructive comments on the draft document.

ISSN 1366-5278

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

M

Contents

List of abbreviations

ACE	angiotensin-converting enzyme
BNF	British National Formulary*
CI	confidence interval
DDD	defined daily dose*
HF	heart failure*
HOPE	Heart Outcomes Prevention Evaluation [clinical trial]
ICD	International Classification of Diseases
MDI	metered-dose inhaler*
MI	myocardial infarction*
NR	not reported consistently*
NS	not statistically significant*
NSAID	non-steroidal anti-inflammatory drug
QALY	quality-adjusted life-year
ROSE	Risperidone Outcome Study of Effectiveness [clinical trial]
SALT	Swedish Aspirin Low-dose Trial
SAPAT	Swedish Angina Pectoris Aspirin Trial
SOLVD	Studies of Left Ventricular Dysfunction [clinical trial]
SSRI	selective serotonin reuptake inhibitor
TD	tardive dyskinesia*

* Used only in tables and figures

Executive summary

Background

Clinical guidelines, defined as 'systematically developed statements to assist both practitioner and patient decisions in specific circumstances', have become an increasingly familiar part of clinical care. Guidelines are viewed as useful tools for making care more consistent and efficient and for closing the gap between what clinicians do and what scientific evidence supports. Interest in clinical guidelines is international and has its origin in issues faced by most healthcare systems: rising healthcare costs; variations in service delivery with the presumption that at least some of this variation stems from inappropriate care; the intrinsic desire of healthcare professionals to offer, and patients to receive, the best care possible. Within the UK, there is ongoing interest in the development of guidelines and a fast-developing clinical-effectiveness agenda within which guidelines figure prominently. Over the last decade, the methods of developing guidelines have steadily improved, moving from solely consensus methods to methods that take explicit account of relevant evidence. However, UK guidelines have tended to focus on issues of effectiveness and have not explicitly considered broader issues, particularly cost. This report describes the methods developed to handle benefit, harm and cost concepts in clinical guidelines. It reports a series of case studies, each describing the development of a clinical guideline; each case study illustrates different issues in incorporating these different types of evidence.

Health economics and clinical guidelines

There has been no widely accepted successful way of incorporating economic considerations into guidelines. Unlike other areas of guideline development, there is little practical or theoretical experience to direct the incorporation of cost issues within clinical guidelines. However, the reasons for considering costs are clearly stated: "health interventions are not free, people are not infinitely rich, and the budgets of [health care] programmes are limited. For every dollar's worth of health care that is consumed, a dollar will be paid. While these payments can be laundered, disguised or hidden, they will not go away"[*]. Such opportunity costs are a universal phenomenon. In the USA it has been recommended that every set of clinical guidelines should include information on the cost implications of the alternative preventive, diagnostic, and management strategies for each clinical situation. The stated rationale was that this information would help potential users to evaluate better the potential consequences of different practices. However, it was acknowledged that "the reality is that this recommendation poses major methodological and practical challenges"[†].

Methods of developing clinical guidelines

A guideline development process summarises the technical information about the value of treatments in a manner that makes them accessible and ready for use in clinical practice, alongside information on contextual issues. The requirement is that the presentation of costs and benefits of treatments is methodologically sound, robust and accessible. This report includes a summary of the current best practice in evidence-based guideline development, including recent methodological advances. The manner in which cost and cost-effectiveness concepts have been successfully incorporated into the guideline process is introduced.

Guideline development case studies

The 'cost-effectiveness' sections of 11 guidelines are reported to illustrate both the range of methods used and the nature of the recommendations

[*] Eddy DM. A manual for assessing health practices and designing practice policies: the explicit approach. Philadelphia: American College of Physicians; 1992.

[†] Institute of Medicine. Guidelines for clinical practice: from development to use. Washington: National Academy Press; 1992.

reached by the guideline development groups when considering the profile of consequences of treatments including costs. These guidelines are broadly grouped as: (1) those using qualitative evidence summary methods; (2) those using quantitative evidence summary methods and addressing relatively narrow clinical questions; (3) those using quantitative evidence summary methods and addressing a broad clinical area; (4) a guideline based upon a decision analysis model.

Conclusions

The focus of this project was to explore the methods of incorporating cost issues within clinical guidelines. However, the process of reviewing evidence in guideline development groups is becoming increasingly sophisticated, not only in considerations of cost but also in review techniques and group process. At the outset of the project it was unclear how narrowly or broadly the concept of 'cost' could be considered. It is now clear that, alongside the effectiveness data and data describing quality of life, cost issues can successfully be represented as part of a profile of treatment attributes. It is also clear that, when used appropriately, modelling processes can provide valuable input into guideline development processes.

Implications of this project

This report describes methods that, in our opinion, are currently optimum for developing clinical guidelines that include consideration of multiple dimensions of evidence (effectiveness, tolerability, harm, quality of life, health-service delivery issues, costs) and it will be relevant to those who commission, develop or use clinical guidelines. The described 'attribute profile' approach to judging whether the costs and consequences of treatments make reasonable sense appears to be the most robust and socially defensible method at this time.

The main implication from this work is that these methods should form the current minimum expected of guideline developers. It is important that the methods described are attempted and developed by other guideline methodologists and health economists and the debate about the valuation of healthcare is expanded.

Recommendations for further research

While working on the case studies a range of unanswered questions were identified, some of which are directly related to the consideration of costs within guidelines and some of which relate to clinical guideline development more generally. Further research should be carried out to answer the following questions.

- What is the relationship between the incorporation of costs into a guideline and the cost impact of a guideline? What are the optimum methods of using cost data in guideline development and of assessing the cost impact of a guideline? Should these processes be unified or separate?
- What are the implications for level of evidence and strength of recommendation taxonomies of considering a range of treatment attributes beyond effectiveness and tolerability?
- What is the role of decision analysis in the development of clinical guidelines?
- In what circumstances is it necessary to use formal consensus methods within a guideline development process?

The research questions above could be usefully informed by the use of more robust designs.

Chapter 1
Introduction

Clinical guidelines, defined as 'systematically developed statements to assist both practitioner and patient decisions in specific circumstances'[1] have become an increasingly familiar part of clinical care. Guidelines are viewed as useful tools for making care more consistent and efficient and for closing the gap between what clinicians do and what scientific evidence supports.[2] Interest in clinical guidelines is international[3,4] and has its origin in issues faced by most healthcare systems: rising healthcare costs; variations in service delivery with the presumption that at least some of this variation stems from inappropriate care; the intrinsic desire of healthcare professionals to offer, and of patients to receive, the best care possible. In the UK, there is ongoing interest in the development of guidelines[5] and a fast-developing clinical-effectiveness agenda[6,7] within which guidelines figure prominently.

The guideline development process recognises the clinician and patient as the decision-making unit, together acting as the arbiter of appropriate treatment. The explicit intent is to attempt to optimise the agency relationship rather than to subvert or bypass it with the mechanistic application of 'evidence'. (The agency relationship in healthcare refers to the observed asymmetry in terms of training, knowledge and experience along with patients' vulnerability, due to illness, that accounts for the considerable influence, desirable or otherwise, that clinicians have on patients' treatment decisions.) An optimised agency relationship would have certain (idealised) characteristics (*Box 1*), while recognising that certain acute or chronic clinical conditions may require delegation of autonomy to carers or the clinician.

There are a number of hypotheses of why healthcare delivery sometimes may fall short of this standard. However, a central problem has been the huge and ever-increasing evidence-base relating to each medical discipline. Modern efforts to summarise this evidence, for example, by the Cochrane Collaboration, have made an important step in making evidence more available. Nonetheless, a systematic review by itself is likely to be insufficient to inform clinical decision-making, since the interplay between evidence from trials and the realities of healthcare delivery must be considered and interpreted (as later examples of guidelines will illustrate).

BOX 1 An optimised agency relationship

- Mandated by the patient, the clinician gathers sufficient data from the patient (history, tests and examination) to be able to make as confident a diagnosis as possible
- The clinician has up-to-date knowledge of the available evidence to interpret diagnostic findings, offer a prognosis and assess the consequences of available patient management options
- The clinician explains to the patient, in appropriate language, the likely consequences of the available management options
- After adequate discussion, with the clinician and carers, the patient chooses their preferred mangement

When advising patients, clinicians balance their own preferences and those of patients and carers, patient-specific information, the benefits, side-effects and safety of treatment and, to varying extents (depending perhaps on the mode of reimbursement), cost. Consequently, the primary goal of guideline development is neither to be content with an odds ratio nor a cost per quality-adjusted life-year (QALY) estimate, since in isolation both are the wrong mode of presentation to inform a doctor–patient interaction. Rather, the objective is to help the clinician and patient to perform an appropriate aggregation of the attributes of treatment, weighing up their relative importance in an individual treatment decision. Clinician and patient are then making appropriate use of the best information available. Implicit in this is that the information is presented in its disaggregated constituent parts (as far as is compatible with addressing the clinical problem). Clinical guidelines have a central role in providing such information.

Over the last decade, the methods of developing guidelines have steadily improved, moving from solely consensus methods to methods that take explicit account of relevant evidence. This improvement should lead to improved guideline validity – and guidelines are valid if 'when

Chapter 2

Health economics and clinical guidelines

Why consider costs?

Unlike other areas of guideline development, there is little practical or theoretical experience to direct the incorporation of cost issues within clinical guidelines. However, the reasons for considering costs are clearly stated by Eddy:[13]

> "Health interventions are not free, people are not infinitely rich, and the budgets of [health care] programmes are limited. For every dollar's worth of health care that is consumed, a dollar will be paid. While these payments can be laundered, disguised or hidden, they will not go away."

Such opportunity costs are not particular to the healthcare system of the USA, but a universal phenomenon; the NHS needs to obtain the best value from finite public funds.

The Committee on Clinical Practice Guidelines[1] recommended that every set of clinical guidelines should include information on the cost implications of the alternative preventive, diagnostic and management strategies for each clinical situation. Its stated rationale was that this information would help potential users to evaluate better the potential consequences of different practices. Although acknowledging "the reality is that this recommendation poses major methodological and practical challenges" it suggested that, in the process of considering costs, five questions should be examined (*Box 2*).

BOX 2 Issues to be addressed in clinical guidelines

- What evidence suggests that the services are likely to affect outcomes for the condition or intervention being considered?
- What groups at risk are most likely to experience benefits or harms from the proposed course of care and its side-effects?
- What is known about the effects of different frequencies, duration, dosages, or other variations in the intensity of the intervention?
- What options in the ways services are organised and provided can affect the benefits, harms and costs of services?
- What benefits, harms and costs can be expected from alternative diagnostic or treatment paths, including watchful waiting or no intervention?

Source: Institute of Medicine, 1992[1]

It then went on to offer a range of reasons why guideline developers would have difficulty finding the answers to these questions (*Box 3*).

BOX 3 Problems confronting guideline developers

- Scientific evidence about benefits and harms is incomplete
- Basic, accurate cost data are scarce for the great majority of clinical conditions and services
- While data on charges may be available, significant analytic steps and assumptions are required to treat charge data as cost data
- Techniques for analysing and projecting costs and cost-effectiveness are complex and only evolving

Source: Institute of Medicine, 1992[1]

An additional layer of complexity is then added by acknowledging that the approach to costs and guidelines will differ between audiences. While accepting that there will be overlap between groups, typically clinicians and patients will be most interested in cost issues impacting on individual treatment decisions. Although this will also be of interest to policy makers, they are likely to have an additional interest in the cost impact of introducing a guideline into a service.

Assessing the cost impact of a guideline

Policy makers may routinely wish to know the cost impact of implementing a guideline. For example, if a new and expensive treatment is recommended, it might be possible to assess the net cost to the NHS of different levels of uptake of recommendations, alongside the expected benefits. However, the longer-term costs and consequences of treatments are often difficult to predict, making overall cost-impact assessments inherently uncertain. Typically, clinical guidelines will cover the breadth of a clinical condition, thereby involving multiple clinical decisions, each with its own associated cost and consequence uncertainties. In addition, the degree of sophistication required to factor in the cost-effectiveness of implementation strategies has seldom, if ever, been addressed. Thus, cost-impact assessments on guidelines may be of less value to policy makers than they suppose.

While cost-impact assessments might be of interest to a guideline development group, their use within the group needs to be handled with care. Cost-impact assessments are potentially problematic since they may be perceived to take the focus away from improving individual treatment decisions and toward healthcare policy traditionally concerned with budgets. Doctors and patients may perceive that 'affordability' rather than 'values' underpins the guideline recommendations, potentially discrediting the guideline medium. However, the distinction is not absolute. Members of guideline groups might perceive a cost-impact assessment to be important at a local level. For example, it may be helpful to know how much a new screening programme will cost, organised at various levels (such as the general practice or primary care group or trust) where delivery issues and implementation are seen as integral to the recommendations. Thus, although assessing the cost impact of a guideline is a legitimate aim, it is not the main subject of the rest of this report, which is concerned with using cost data when developing a clinical guideline.

Using cost data when developing a clinical guideline

Questions about, and limitations of, cost-effectiveness analyses raise the issue of how to use cost data in a guidelines group. Should data be presented alongside recommendations based solely on clinical effectiveness or incorporated into the judgement process of deriving recommendations? Williams[14] and Eddy[13] argue that guidelines based on effectiveness issues and then costed may differ substantially from, and be less efficient than, guidelines based on cost-effectiveness issues. The complexity of this process, and the reactions it evokes, are reflected by the Committee on Clinical Practice Guidelines' report of "much debate, and with some vigorous dissent".[1] There has been no widely accepted successful way of incorporating economic considerations into guidelines.

The measurement of health has mushroomed as an academic and clinical pursuit in recent years and generic measures have been developed that express patient health status (and its changes over time) as a single index (e.g. the QALY). Thus, theoretically, health gains could be compared across different diseases and patient groups, and cost–QALY estimates could provide a common metric for comparing the value for money of the myriad healthcare interventions available. To maximise health in a population would require the various cost-effective strategies across all therapeutic areas to be implemented in proportions that achieve the socially optimal allocation of resources. This would mean ranking interventions in order of cost-effectiveness and then working down the list until the budget is spent. Such a ranking would thus define the range of available treatment options and clinical guidelines would only need to consider issues of effectiveness and harm. Unfortunately, such ranking is not possible.[15] The number and variety of interventions, patients, settings and other variables is huge, and our knowledge of the value of treatments sparse, making this an impractical option.

The rationale underpinning such methods has been the belief that complex cost and benefit profiles associated with a range of treatments can be aggregated, producing 'an answer' to aid decision-making (at least with respect to efficiency). This has proved unproductive, in part because the methods and data have been inadequate for the provision of a simple answer and in part because clinicians (the key audience) do not appear to think of appropriate healthcare in terms of economic outcomes such as cost-effectiveness ratios.

The fact that there is an issue about how healthcare professionals think about, and react to, explicit cost issues in guidelines is understandable. Most healthcare professionals have a limited knowledge of health economics and economic modelling. Guidelines based on clinical effectiveness could be enhanced or undermined by the incorporation of economic considerations, depending on whether they are seen as attempts to achieve cost-effectiveness or cost-containment. It remains a research issue as to how the incorporation of economic considerations will affect the use of guidelines in individual treatment decisions, although the intention is to encourage a more explicit consideration of costs and consequences in each consultation at which guidelines are used. British healthcare professionals are not accustomed to this process at anything other than an implicit level, although in recent years cost messages have indirectly impinged more and more through formulary lists or fund-holding initiatives. In the absence of an overarching allocative framework (and in a system in which healthcare is provided from general taxation) the first step in moving this process forward is to develop robust methods of incorporating economic issues into clinical guidelines.

Chapter 3

Methods of developing clinical guidelines

Background

This chapter summarises the current best practice in evidence-based guideline development, including recent methodological advances. The manner in which cost and cost-effectiveness concepts have been successfully incorporated into the guideline process is introduced. A guideline development process summarises the technical information about the value of treatments in a form that makes it accessible and ready for use in clinical practice, alongside information on contextual issues. The requirement is that the presentation of costs and benefits of treatments are methodologically sound, robust and accessible. The novel aspect is the dynamic development and use of economic data (rather than the use of static data from published studies), alongside traditional clinical inputs, in the development of clinician valuation of treatments and consequent guideline recommendations.

Methods of developing clinical guidelines

There are common elements in the methods of guideline development described in North America[1,13] and in the UK;[9,16–20] these are summarised in *Box 4*.

BOX 4 Five steps in clinical guideline development

1. Identifying and refining the subject area of a guideline
2. Convening and running guideline development groups
3. Obtaining and assessing the evidence about the clinical question
4. Translating the evidence into a clinical guideline
5. Arranging external review of the guideline

In recent years, each step in the guideline process has developed methodologically and it is appropriate to summarise the current state of the art in this process, showing how economic concepts have become incorporated.

Identifying and refining the subject area of a guideline

Potential areas for which guidelines could be developed can emerge from an assessment of the major causes of morbidity and mortality for a given population, uncertainty about the appropriateness of healthcare processes, poor or variable uptake of new and worthwhile interventions, or the need to conserve resources in providing care. Within any given area, guidelines may address narrow or broad questions and be condition-based (e.g. available treatments for diabetes or coronary artery disease) or procedure-based (e.g. alternative approaches to hysterectomy or coronary artery bypass surgery). Given the large number of potential areas, some form of prioritisation is needed to select a particular question or focus for guideline development.

Refining the subject area of a guideline

In whatever way the topic for guideline development is initially identified, it will usually need to be refined before an assessment of the evidence is begun, in order to answer exact questions. This can be achieved in a number of ways. The usual method is a dialogue among clinicians, patients, and the potential end-users or evaluators of the guideline. This will normally be conducted before guideline development begins but will often continue as discussions around the emerging evidence take place within the guideline development panel.

Failure to carry out this refinement runs the risk of leaving too broad a scope in the clinical condition or question. For example, a guideline on 'the management of diabetes' could cover primary, secondary and tertiary care and multiple aspects of management, such as screening, diagnosis, dietary management, drug therapy, risk-factor management and indications for referral from primary to secondary care. All of these could be legitimate areas for guideline development, but the task of developing a guideline covering all of them would be considerable. Therefore, a group needs to be clear about the areas within the scope of their activities. It is possible to develop guidelines that are both broad in scope and

evidence-based, but to do so usually requires a considerable investment in time and money, both of which are easy to underestimate.

Two methods to help define the clinical question of interest and identify care processes for which analysis of available evidence is required are the construction of disease pathways and causal pathways.[21] A disease pathway approach seeks to identify and quantify the common paths of patients as a disease runs its course and to identify key intervention points. A causal pathway illustrates, in the form of a diagram, the linkages between intervention(s) of interest and the intermediate, surrogate, and health outcomes that the interventions are thought to influence. In designing the pathway, guideline developers make explicit the premises on which their assumptions of effectiveness are based and the outcomes (benefits and harms) that they consider important. This identifies the specific questions that must be answered by the evidence in order to justify conclusions of effectiveness and also highlights gaps in the evidence for which future research is needed.

Epidemiological summary

To provide context to the review of treatments, it is useful to make a presentation of the epidemiological profile of the disease and current patterns of care. This frames the tasks set for the guideline development group, providing an understanding of prognosis, morbidity and mortality as well as health-service resource use. It is valuable to have in view the size of the problem being considered and its effects upon the lives of patients and carers. An added benefit is the descriptive use of cost data, categorising the current use of resources. This provides an opportunity (non-threatening for those diffident about economic concepts) for the group to discuss whether it may be possible to make better use of resources and so to introduce value-for-money concepts. To date, such summaries have not been available at the scoping stage of guideline development. However, when available, an epidemiological summary can be an important element of the process of refining the subject area of the guideline.

Guideline development groups

Convening and running guideline development groups

There is no single right way to organise guideline development groups. The amount of work involved is often considerable and it is important to ensure that guideline development is adequately resourced. Groups often report underestimating the resources required for the task – not only in terms of finance but also in project management and general administrative support. Recent North of England guidelines[19,22–30] have utilised a specialist resource team inside the guideline development group to undertake identification, synthesis and initial interpretation of relevant evidence, the convening and running of the guideline development groups, and the production of the resulting guidelines. The full guideline development group has then been responsible for discussing the implications of the evidence and drawing up appropriate recommendations.

The guideline development group: membership and roles

The composition of a guideline development group can be considered in two ways: by the disciplines or affiliations of the group members who are stakeholders in the area of the guideline; and by the roles required within the group.

Disciplines or affiliations of group members

Identifying stakeholders involves identifying all the groups whose activities are covered by the guideline or who have other legitimate reasons for having input into the process. This is important to ensure adequate discussion of the evidence (or its absence) when developing the recommendations in the guideline. The need for such a multidisciplinary approach is borne out by empirical evidence showing that, when presented with the same evidence, a single speciality group will reach different conclusions from a multidisciplinary group, with the former being systematically biased in favour of performing procedures in which the speciality has a vested interest.[31–33] For example, the conclusions of a group of vascular surgeons favoured the use of carotid endarterectomy more than a mixed group of surgeons and medical specialists.[33] There are good theoretical reasons to believe that individuals' biases are better balanced in multidisciplinary groups, and that such balance will produce more valid guidelines. Ideally the group should have at least six but no more than 12 to 15 members: having too few members limits adequate discussion and having too many members creates difficulty with the effective functioning of the group. Under certain circumstances (e.g. when developing guidelines for broad clinical areas) it may be necessary to trade off full representation of all possible stakeholders against the requirement to have a functional group.

Roles

The roles required within guideline development groups are: group member; group leader; specialist resource; technical support; and administrative support. Group members, as indicated above, are invited to participate as individuals working in their field. Their role is to develop recommendations for practice, based upon the available evidence and their knowledge and experience.

The role of the group leader is to ensure both that the group functions effectively (the group process) and that it achieves its aims (the group task). Although guideline development groups are often chaired by pre-eminent clinical experts in the topic area, it is possible that the process is best moderated by someone familiar with (though not necessarily an expert in) the management of the clinical condition and the scientific literature but who is not an advocate.[9] Such an individual acts to stimulate discussion and allow the group to identify where true agreement exists, but does not inject personal opinions into the process. This role requires someone with both clinical skills and group process skills. There is also evidence that conducting the group meetings using formal group processes rather than informal ones produces different, and possibly better, outcomes.[34–37]

Guideline processes will require the support of a variety of specialists. Some of the potential skills required are shown in *Box 5*. Finally, groups will need administrative support for such tasks as preparing papers for meetings, taking notes and arranging venues.

BOX 5 Skills needed in guideline development

- Literature searching and retrieval
- Epidemiology
- Biostatistics
- Health-services research
- Health economics
- Clinical expertise
- Group process expertise
- Writing and editing

The guideline development group explores, within the clinical area of the guideline, all of the situations for which there may be a need to offer recommendations. Although the topic has been pre-defined, in the first meeting the group is asked to confirm their acceptance of both the clinical content of the areas of the guideline and the scope of the questions to be answered within it. This ensures a shared view of group aims between the group and the research team and enables the group leader to challenge deviation from the task in hand. From a practical viewpoint, a review of the evidence needs to begin before the first meeting. However, the group has the option of extending, restricting or refining the scope of the review. Given the amount of work often involved, decisions to alter the scope of work should remain centred upon the value of subsequent information in deriving treatment recommendations.

Consumer representation

Finding appropriate means of taking patient views into account is recognised to be an unresolved methodological issue.[38–40] In common with others, the North of England guideline development group had had mixed experience when trying to include a 'token' patient representative. One approach may be to assume that the objective of guidelines is to inform the doctor–patient relationship about the best available evidence on the various attributes of treatment, at which point, patient-specific information must always be introduced into the decision-making process.[41] This may be a valid stance when considering the role of individual drugs, but may be less supportable when considering packages of care for 'whole' diseases. However, in these circumstances, rather than direct involvement in the group, it may be more productive for patients to contribute to the content of a guideline through carefully constituted focus groups, or some similar forum. This would entail patients (or their representatives) exploring the available evidence with a trained facilitator, with comments fed back to the guideline process.[38,40] This type of activity is outside the scope of work presented in this report, but may be adopted in ongoing national guidelines.

Assessing the evidence

The core of any guideline is the systematic review of evidence to lead the group in an informed debate about the value of treatment alternatives. Given the increasing availability of systematic reviews, we give separate consideration to the evidence from primary and secondary research.

Evidence from primary research

Identifying evidence: the search strategy

The aim is to identify and synthesise relevant published and unpublished evidence to allow

recommendations to be evidence-based wherever possible. A sensitive search should be carried out using the electronic databases (e.g. MEDLINE, EMBASE, SIGLE and the Cochrane Controlled Trials Register), attempting to locate systematic reviews and meta-analyses, randomised controlled trials, quality-of-life studies and economic studies using a combination of subject heading and free text searches. When retrieving studies, extensive use can be made of high-quality recent review articles and bibliographies. Experts in the subject area can also be contacted. It may be efficient to update existing systematic reviews when these are unable to provide valid or up-to-date answers. The process of literature searching has become highly technical and is best conducted or supported by trained personnel. The expert knowledge and experience of group members backs up the search strategy.

Assessing and synthesising the literature

Retrieved studies are assessed for their quality, concentrating on questions of internal validity (the extent to which the study measured what it was intended to measure), external validity (the extent to which study findings could be generalised to other treatment settings) and construct validity (the extent to which measurement corresponded to theoretical understanding of a disease).[42] The specific dimensions of quality examined in each study are reported in *Box 6*.[19]

BOX 6 Quality criteria for the internal validity of randomised trials

- Appropriateness of inclusion and exclusion criteria
- Concealment of allocation
- Blinding of patients
- Blinding of health professionals
- Objective/blind method of data collection
- Valid/blind method of data analysis
- Completeness and length of follow-up
- Appropriateness of outcome measures
- Statistical power of results

Once individual papers have been checked for methodological rigour and clinical significance, data are extracted on the benefits, harms and, where applicable, resource implications of the interventions being considered. These are usually presented in a form that facilitates easy comparison of the designs and results of studies. This can be done in two ways: qualitatively or quantitatively. Qualitative or narrative methods may be necessary when relevant evidence is heterogeneous in terms of study design or outcomes.

Papers are categorised according to study design, reflecting susceptibility to bias. An example of an evidence categorisation is shown in *Box 7*. This is adapted from the US Agency for Healthcare Policy and Research Classification.[43] This categorisation is most appropriate to questions of causal relationships.

BOX 7 Categories of evidence

Ia	Evidence from meta-analysis of randomised controlled trials
Ib	Evidence from at least one randomised controlled trial
IIa	Evidence from at least one controlled study without randomisation
IIb	Evidence from at least one other type of quasi-experimental study
III	Evidence from non-experimental descriptive studies, such as comparative studies, correlation studies and case–control studies
IV	Evidence from expert committee reports or opinions and/or clinical experience of respected authorities

Some health-service activities (e.g. prognosis or diagnosis) may not be best evaluated using the randomised controlled trial approach and other evidence-grading taxonomies may be appropriate in these instances.

Questions are answered using the best evidence available. When considering a question of the effect of an intervention, if the question can be answered by evidence provided by a meta-analysis or randomised controlled trial, then studies of weaker design (controlled studies without randomisation) are not reviewed. If the evidence summary is quantitative, then where studies are of poor quality, or contain patient groups that are considered *a priori* to be likely to have different responses, the effects of inclusion or exclusion are examined in sensitivity analyses. If data on relevant outcomes are missing from studies and unavailable from authors or sponsors, these studies cannot be included in meta-analysis, thus introducing a form of publication bias. The impact on subsequent treatment recommendations of any such exclusion needs careful consideration.

Evidence from secondary research

In some circumstances it may be possible to use previously conducted systematic reviews and this may be a necessity if there are inadequate resources to conduct a new review. Whatever the circumstances, previously conducted reviews need to be carefully assessed for the quality of their methods and presentation of findings.

Systematic reviews used in guideline development can be considered, as with the primary studies, in terms of their internal and external validity. For a review, internal validity relates to whether or not the review is offering precise summary measurement of what it purports to measure. External validity then relates to the degree to which the findings of the review can be generalised – in this case, to the healthcare setting considered within the guideline. When using existing systematic reviews, there are potential problems with the summary metric used and the possible need to update a review.

Internal validity

Internal validity relates to the identification of the original studies included in the review and the method of conducting the review. As a result of publication bias or failure to find relevant studies, reviews of selected studies can reach biased conclusions[44,45] and should not be used without further work. Similarly, those reviews using inappropriate or flawed methods cannot be taken at face value and are probably best excluded.

When judging the internal validity of a review, a guideline development group will also have to decide whether or not the review addresses all of the dimensions on which evidence is needed to derive recommendations. Many systematic reviews are concerned with obtaining summary measures of effectiveness and do not extend to cover other issues of concern to patients such as side-effects, tolerability, or consequences for work-place activities. While the restricted focus may be a constraint of the studies contributing to the review, this is not necessarily the case. Failure to consider significant side-effects may offer an inappropriately positive view of an intervention and limit the validity of a review. Similarly, resource implications and costs, not normally the subject of systematic reviews, are likely to be an important consideration in the implementation of a guideline.

External validity

The external validity of a review is more likely to be problematic, although it is important to recognise that this may again be as much to do with the original studies as with the review. A key aspect of a review is the rationale for including or excluding studies. As with primary studies, the generalisability of a review will be limited by the characteristics of the populations included in it and the settings in which the studies were conducted. Ideally, a guideline development group wants evidence from studies that include people who are typical of those addressed by the guideline. Therefore, the minimum requirement is that a review lists the important characteristics of the population within each study. However, this will not provide information on other aspects, such as recruitment rates in studies. The setting in which a study is conducted will also influence its external validity, particularly if it relates to elements of service delivery. The exclusion criteria for studies used in a review may be acceptable (for example, a trial is not randomised when other available and adequate trials are) or debatable (for example, using a different dose of drug from the one used in practice) and merit careful consideration. Often there may not be a clear-cut answer. A guideline development group wanting to know about the effectiveness of education delivered in a primary care setting to patients with asthma may question a review including studies from secondary care settings. In such instances the studies have to be reviewed to identify their applicability to health-service delivery in primary care.

There are two further ways in which the generalisability of a previously published review may be undermined.

Firstly, it is unlikely that a review can address all the questions posed within a guideline development process. Relevant clinical questions are defined by the guideline development group with the aim of deriving recommendations that can appropriately inform doctor–patient interactions: there may be a substantial contextual component to decision-making. Although well-conducted systematic reviews may be available, they are unlikely to address all the issues of interest to a guideline development group. Indeed, it would be surprising if a review conducted outside a guideline development process could second-guess all the relevant questions.

Secondly, in a review that uses meta-analysis as a method of summary, the process will require common estimates of effect that can be summarised across studies. Particularly in complex

clinical areas where multiple and differing outcome measures have been used within studies, the choice of studies to include in a review may be influenced by factors such as whether or not common end-points are available, rather than by the clinical usefulness of the studies.

Summary metric used

The odds ratio is the most frequently used and statistically most robust metric for summarising effectiveness within systematic reviews. However, the odds ratio alone is insufficient to summarise the evidence from trials. Used alone, it is not readily interpretable and it needs to be considered alongside a summary statistic of absolute risk reduction.[46] A systematic review that presents the odds ratio alone is difficult, if not impossible, to use. The only option would be to assume that the odds ratio could be applied to the level of baseline risk in the population within which the guideline would subsequently be used. For conditions for which the level of baseline risk might be available, this strategy involves a number of assumptions. However, for many conditions for which one might want to develop guidelines such information is not available.

Obtaining an updated review

If a relevant systematic review is identified within a guideline development process, there may be an issue about its timeliness. In a clinical area within which a review was completed 2 years ago, there may well have been subsequent relevant papers published that have not been incorporated into the review. Under such circumstances, the ideal is for the review authors to be contacted for an update. If this is not possible, then a guideline process has to either replicate the whole review or consider the new papers alongside, but outside, the review. Under such circumstances, conflicting results may be difficult, if not impossible, to reconcile.

Describing the value of healthcare interventions

A key aspect of the functioning of a guideline development group relates to the manner in which summary evidence from trials is presented and interpreted. Superficially, a trial finding presented as an odds ratio may appear far more impressive than one presented as an absolute risk reduction: each measure has a different meaning and different strengths and weaknesses. Time must be invested in the group to ensure consistent and informed interpretation of trial findings. A brief summary of the various measures available illustrates aspects of group training and the way measures have been developed and used in guideline development groups to interpret evidence.[46]

Binary outcomes

Meta-analysis of binary data, such as the number of deaths in a randomised trial, enables the results of a group of trials to be expressed in a number of ways. These are primarily odds ratios, risk ratios (also known as relative risks) and risk differences. If binary data on mortality from a trial are expressed in a 2 × 2 table:

	Dead	Alive
Intervention group	A	B
Control	C	D

odds ratios are defined as: $\frac{A}{B} \Big/ \frac{C}{D}$

In other words, the odds ratio is the odds of death in the intervention group (number of deaths divided by the number of survivors) divided by the odds of death in the control group.

Risk ratios are defined as: $\frac{A}{A+B} \Big/ \frac{C}{C+D}$

The risk ratio is the risk of death in the intervention group (number of deaths in the intervention group divided by the total number allocated to the intervention) divided by the risk of death in the control group. Trials sometimes refer to relative risk reductions, which are calculated as (1 – risk ratio).

Risk differences are defined as: $\frac{A}{A+B} - \frac{C}{C+D}$

The risk difference (also known as the absolute risk reduction) is the risk of death in the intervention group (number of deaths in the intervention group divided by the total number allocated to the intervention) minus the risk of death in the control group. Number needed to treat, an alternative presentation, is calculated as (1/risk difference).

Worked example

In a trial of an angiotensin-converting enzyme (ACE) inhibitor in patients with heart failure, at end-point there were 452 deaths among 1285 patients randomised to receive enalapril,

and 510 deaths among 1284 patients allocated to control treatment.[47] In a 2 × 2 table this is:

Treatment group	Dead	Alive
Intervention	452	833
Control	510	774

providing an odds ratio of 0.82, a risk ratio of 0.89, and a risk difference of –0.045 (or a 4.5% reduction in the risk of death).

The odds ratio is a statistically robust measure, but is hard to interpret clinically. It may be particularly useful when attempting to combine studies that are estimating the same common underlying effect, but among which both severity of condition and length of follow-up may differ substantially. This occurs because the odds ratios express the relationship between rates, rather than those rates in absolute terms. However, the odds ratio is insufficient for clinical decision-making alone: an intervention with an impressive odds ratio will not lead to large benefits in practice if the events are rare; conversely, an intervention with a small odds ratio may have a substantial impact if events are very common.

Risk differences are not very helpful for exploring common underlying effects, but are very useful for describing the practical importance of the effects of an intervention in practice. A standard problem with the risk difference is that it is often derived from trials that have differing lengths of follow-up. One of the main potential advantages of the risk difference is that it enables the practical value of interventions to be assessed and compared with alternative treatment strategies. Thus the incidence risk difference is used to estimate treatment effects across a common time frame, for example the number of deaths avoided as a result of treating 1000 patients for a year.[19]

A **confidence interval** (CI) for a treatment effect estimated in a trial is the interval in which the underlying population treatment effect is assumed to lie, with a specified probability. The specified probability is arbitrary: 95% is the most commonly chosen value, meaning that the true underlying treatment effect is assumed to lie within the interval 19 times out of 20. In figures describing groups of studies combined by meta-analysis, the point estimate of effect from each study is indicated, as are the 95% CIs, which are denoted by horizontal lines: the shorter the line, the narrower the CIs and the greater the precision of measurement in the study. The best and most likely estimate of effect is the point estimate at the centre of the CI range.

Binary versus continuous outcomes

Often, in randomised clinical trials the effectiveness of treatments can best be expressed by a binary outcome: for example, alive or dead. In these cases odds ratios are useful for describing underlying treatment effects, and risk differences or numbers needed to treat are useful for describing the absolute size of the effect of treatment in the trial populations. However, many outcomes are not amenable to this binary approach, and are better considered as continuous measures.

The approach of dichotomising data that are naturally continuous (for example into treatment failures and successes) is to be discouraged. It is often arbitrary, and may result in pooling different scores based on different cut-offs in different studies or cut-offs that have been identified with knowledge of the data and thus show the data in a particular light. The approach may exaggerate small differences in effect and, more fundamentally, it explains the data poorly.

Meta-analysis of continuous data

Where studies use a common outcome measure, meta-analysis can combine these to calculate a summary weighted mean difference comparing treatment and control groups. When there are concerns that there are differences between studies in the metric used, standardised scores are calculated for each trial. Examples might be where different but related instruments have been used to estimate the same common underlying effect in patients with schizophrenia, or when there is likelihood of poor inter-rater reliability in the use of instruments. In such an event the approach advocated by Hedges and Olkin,[48] in which the standard deviation for each study is based upon a weighted mean of the intervention and control group variances, has been used in guidelines.

Trial phases

Double-blind randomised trials are occasionally criticised for inadequately representing treatment in the real world. In other words, trials that use a well-defined population without co-morbidity, limit treatment options and make both the doctor and patient blind to treatment may provide different results from those realised in practice. The evaluation of pharmaceuticals is

best undertaken using a series of experimental studies. This is reflected in Phase II and Phase III studies (small-scale dose ranging and licensing). For studies in Phase IV some of the requirements of the earlier trials may be relaxed to better reflect the real world: these less stringent requirements may include relaxation of blinding and of limitations on clinical strategies such as choice of drug after initial randomisation, and inclusion of patients with co-morbidity. Such studies have been described as 'contaminated with the real world'[49] and it may be difficult to work out what is being estimated (particularly with, say, strong patient or doctor preferences for one treatment). However, when examined with the earlier Phase III trials, Phase IV studies may add useful information.

Meta-regression analysis

Where a number of trials examine the same underlying question, techniques may be used to construct regression models to provide the best estimate of the predictive value of a factor.[50] For example, this approach could be used to provide best estimates of the predictive value of whether cholesterol level predicts outcome in trials of statins.

Estimating costs

While a social perspective in economic evaluation is both desirable and formally correct, in practice, because of the (un)availability of data, analyses of cost are often limited to costs borne by the NHS. The best sources of data for resource usage and cost are a controversial matter. Resource data from trials may be artificial, whereas observational health-service data comparing the costs of treatments may suffer from a range of biases. The need for meta-analysis of clinical end-points stems from the fact that alternative treatment strategies often feature small differences in outcome. Consequently, precise and internally valid trials are required to achieve a reliable and precise measure not just of differences in health outcome but also of (correlated) differences in resource consequences. The assumption is that while absolute use of resources in trials may be atypical, the difference between treatments will be less so. The approach adopted by the guidelines group is to apply the same categories of evidence used for effectiveness to resource use and to establish the generalisability and relevance of findings by mapping their consequences onto current national patterns of resource use.

Observational data (e.g. insurance claims databases) on resource use should not be used to explore differences between interventions because of the unknowable biases in these data. However, observational data have been used to validate the predictions arising from review profiles and models.

Unit cost data used in guidelines are in the public domain; it is beyond the scope of the guideline development process to conduct new costing studies. Costs can be calculated by attaching published average national unit costs to resource items. Economists often argue that, for decision-making purposes, marginal costs are preferable to average costs.[51] While the problems associated with average costs are recognised, there is no generally valid or accepted method for presenting marginal costs on items or procedures: these will vary from locality to locality. The simple presentation of analyses permits decision-makers to apply different unit costs when such information is locally available. Reflating of unit costs from different years of origin to a common year, to adjust for healthcare cost changes over time, can be carried out. However, reflating was not used in the guidelines summarised in this report because, in general, there was no more than a 2-year gap between the oldest and newest values. Additionally, reflating is an ambiguous practice for certain items such as drug costs for which, under UK reimbursement, the price tends to remain fixed over substantial periods of time.

Building a profile of treatments

The product of the review of evidence, for a guideline development group, should be an appropriate summary or profile of the important consequences of treatment. This may include evidence about clinical outcomes, compliance, quality of life, safety, and health-service resource use (*Box 8*).

BOX 8 The profile of treatment attributes addressed in guidelines

- Effectiveness
- Quality of life
- Tolerability
- Safety
- Health-service delivery issues (implementation)
- Health-service resource use
- Health-service costs
- Patient and carer costs

Any aspect of treatment that may be valued by patients or society should be considered for inclusion in the profile. Dogmatic adherence to the list of attributes is unnecessary, since different diseases and their treatments will impact upon patient health differently and anyway data are often unavailable for a number of attributes. However, it is important, early in the guideline process, to identify the important consequences of alternative strategies of care.

The profile approach leads to simple presentations in the guidelines of cost implications and other consequences that are readily comprehensible to readers of any background. The available evidence on which these presentations are based is not necessarily robust, but by explicitly identifying uncertainties, the presentation of the evidence accurately identifies strengths and weaknesses, and end users of the guideline can easily explore alternative values.

Substantial use of meta-analysis may be possible when the data permits. However, summaries of the evidence may be limited to a narrative-style review (a qualitative summary) when available evidence is poor, uses incompatible assessments of outcome, or is inconsistently reported. Sometimes the information emerging from the review profile does not provide a clear message and it may be useful to use some form of modelling to help the guideline development group to explore the implications of treatments more fully.

The review, and need for any subsequent modelling, evolves through group discussion as understanding of the value of treatments emerges.[52] The group uses the profile to explore the incremental costs and consequences of the different healthcare decisions open to them to recommend. Economic analysis is thus used to attempt a robust presentation showing the possible bounds of cost-effectiveness that may result. The range of values used to generate low and high cost-effectiveness estimates reflects the available evidence and the concerns of the guideline development group. Nonetheless, the simplicity of presentation permits simple rework-ing with different values from the ones used. Recommendations are graded to reflect the certainty with which the costs and consequences of a medical intervention can be assessed. This practice reflects the desire of group members to have simple, understandable and robust information based on good data.

There may be evidence that health-associated costs borne by patients and carers (e.g. travel and time to receive care, over-the-counter drugs, disability costs) and indirect costs of lost earnings differ significantly between alternative treatments. This should be considered relevant to a treatment decision at least in as much as it may undesirably influence compliance with treatment. There is the possibility that organisational alternatives may shift costs from the health service to individuals and the appropriateness of this may depend on the disease considered and contextual circumstances. Seldom are there adequate data to address costs borne by patients, but where this is a concern these costs can be described as attributes of treatments.

In many fields of healthcare there is a body of economic literature accompanying the clinical studies. As with a review of available clinical trials, it is feasible to have a summary of published economic analyses, and methods of assessing and categorising economic studies are being developed. Unlike protocol-driven prospective clinical trials, economic analyses are usually retrospective and the analyst has the choice of how to construct the model and use the data. Qualitatively, there is far greater scope for bias, either explicit or implicit, in the process or model construction, reporting of findings and exploration of uncertainty.[53,54] A guideline development process should lead to the best available presentation of the known costs and various physical consequences of treatment alternatives. These data are unlikely to match the baseline assumptions in any published model. Each clinical trial presents unique or independent data and the trials together can be summarised to obtain an overview. Different economic analyses take different cuts at the same clinical data and there is no quantitative way to summarise the findings of all of the analyses because the data are not independent: thus, there is no 'weight of evidence'. Published decision analyses are often not transparent and it can be difficult and time-consuming to validate the findings presented. In some areas of medicine, a thorough review of published economic analyses would be a mammoth task with little obvious return.

Sometimes one or two economic analyses may have had considerable influence upon a clinical field. In such circumstances it may be useful to summarise these as part of a guideline development process and comment on their findings compared with those of the guideline.

Translating the evidence into a clinical guideline

Groups reach their decisions about the interpretation of the evidence set before them by consensus. Traditionally this has been an informal process, albeit structured by the evidence. The evidence for treatment alternatives is summarised in a series of graded statements and these are used to formulate treatment recommendations. While formal decision-making processes within guideline development have been described,[35,37] their use is still largely experimental with little indication of when it is important. In some guideline areas there will be no valid evidence and recommendations will either have to be consensus-derived and based solely on opinion or not derived at all. When the process identifies important unanswered questions, these should be recorded as a research item.

Strength of recommendation

Treatment recommendations are graded A to D as shown in *Box 9*. The guideline distinguishes between the category of evidence and the strength of the associated recommendation. It is possible to have methodologically sound (category I) evidence about an area of practice that is clinically irrelevant or has such a small effect that it is of little practical importance and would attract a lower strength of recommendation. More commonly, a statement of evidence would only cover one part of an area in which a recommendation has to be made or would cover it in a way that conflicts with other evidence. To produce comprehensive recommendations, the group has to extrapolate from the available evidence. This may lead to weaker levels of recommendation (B, C or D) based upon evidence category I statements. It is not assumed that guideline group members will always be able to reach agreement on the interpretation of research evidence, and the nature of such disagreement is reflected in the text of the guideline.

BOX 9 Strength of recommendations

A	Directly based on category I evidence*
B	Directly based on category II evidence **or** extrapolated recommendation from category I evidence
C	Directly based on category III evidence **or** extrapolated recommendation from category I or II evidence
D	Directly based on category IV evidence **or** extrapolated recommendation from category I, II or III evidence

* *See* Box 7, *page 8*

It is technically possible to have a transparently worthwhile treatment that has no experimental evidence to support it: an example might be the decision to immobilise a fractured bone to produce pain relief and allow healing. Consequently the Royal College of General Practitioners' diabetes guideline methodological group decided to allow the potential for upgrading of self-evident but inadequately evaluated interventions to receive higher grades of recommendation. Such freedom is potentially useful but not without risks: group members new to systematic enquiry may think 'strong personal belief' and 'self-evident benefit' to be synonymous.

The working of the group can be characterised as a dynamic process in which an understanding of the pros and cons of treatment emerges and is refined, questions are responded to with available evidence and uncertainties are assessed.[19] The grades attached to recommendations are determined by the overall quality of evidence as interpreted by the group. The use of a broad range of attributes of treatment allows the basis for making recommendations to be precisely defined (*Box 10*).

BOX 10 Factors contributing to the process of deriving recommendations

- The nature of the evidence (e.g. its susceptibility to bias)
- The applicability of the evidence to the population of interest (its generalisability)
- Resource implications and their cost
- Knowledge of the healthcare system
- Beliefs and values of the panel

Guideline documentation

The 'optimal' format of a guideline report is unknown and, for many, arbitrary: relatively few users of the findings will want to reference the 'full work' but will receive summaries of varying degrees of detail. Nonetheless it is important that such summaries are supported by an accessible resource document (*Box 11*).

To reflect their rationale and their intended use to guide, all guidelines should carry a clinical disclaimer that guideline recommendations do not replace the clinician's obligation to practice safe and appropriate medicine. This disclaimer has been developed to include an 'economic' disclaimer: 'decisions to adopt any particular recommendation must be made by the practitioner in the light of available

BOX 11 Aspects of a guideline resource document

- Group membership and conflicts of interest
- Responsibilities of individuals
- Funding
- Methods of reviewing and group process
- Clear objectives for the guideline and explicit criteria for assessing these
- Limitations of scope and process, and disclaimers
- Descriptions of evidence about treatment alternatives
- Accurate referencing of evidence
- Graded summary statements of evidence
- Linked graded treatment recommendations
- List of external reviewers
- Expiry date and contact point for guideline updates

resources and circumstances presented by individual patients'.

External review

It is invaluable to have a peer review of a provisionally completed guideline. The input of reviewers in the guideline documentation should be acknowledged. Although the guideline content remains the responsibility of the development group, external feedback by subject-area experts provides important quality assurance and may identify inadequacies or flaws. Reviewers should provide statistical, health-service research and economic expertise as well as disease-specific expertise.

Surveillance and updating of recommendations

It is important to know when treatment recommendations were derived and whether, because of the availability of new evidence, recommendations may have passed their 'sell-by date'. Sometimes guideline developers know that important trials are on the horizon, but predicting when treatment recommendations may cease to apply is often arbitrary. Guideline developers should provide a contact point so that enquiries about updates can be fielded, and any decision rules that may lead to reconvening of a guideline development group should be specified (if known).

Summarising the guideline development process

Guideline development methods have developed dramatically from their roots in consensus approaches to incorporate evidence-based medicine and cost-effectiveness concepts. The intention is to provide important and transparent treatment recommendations to busy health professionals. However, guidelines have proliferated to such an extent and use such a wide range of standards and practices that they may be self-defeating. Health professionals do not have the time or expertise to assess the validity of the guidelines, and there is a clear need to clarify the use of guidelines in the NHS. A summary of aspects of the guideline development process discussed in this chapter is shown in *Box 12*.

BOX 12 Aspects of a guideline development process*

- Identify and refine the subject area of a guideline
- Conduct background scoping work on treatments
- Obtain adequate resources and skills to address the subject area
- Define a protocol and conduct a systematic review
- Convene a guideline development group with appropriate representation

In the group

- Conduct appropriate training on the guideline process and the interpretation of systematic reviews, meta-analysis and economic analysis
- Discuss current understanding of the disease, epidemiology and health-service resource use
- Present and discuss systematic reviews addressing the subject area
- Derive statements about the strength of evidence in the review
- Develop a profile of the costs and consequences of treatment options
- Augment the profile with modelling exercises if this will help the group to explore the value of treatments
- Derive and grade treatment recommendations
- Identify an expiry date for the guideline recommendations
- Draft and circulate a full guideline report for comments

At the end of the process

- Submit a full guideline report to external referees
- Reconvene the guideline development group if important omissions are identified

* *Most tasks are ongoing processes: sequencing of activities may vary*

Chapter 4

Case studies of guidelines that use qualitative evidence summary methods

Three guidelines were developed using qualitative summary methods. They focused on issues of effectiveness and because of the paucity of adequate comparative data included only brief consideration of the issues of potential harm, quality of life or cost.

The primary care management of stable angina

The aim of this guideline[55] was to provide recommendations to aid primary care health professionals in the symptomatic management of patients with chronic stable angina. The guideline development group reconvened in 1999 to update the first version of the North of England Stable Angina Guideline produced in 1996.[56] The areas covered by the guideline were: investigation, risk-factor identification and management, drug treatment, and referral. Unstable angina or myocardial infarction were not covered.

The cost-effectiveness of primary care management of patients with angina

Patients presenting with angina in general practice are likely, over time, to be high users of healthcare resources. Ideally, it would be desirable to know how different strategies of diagnosis, management and use of drugs, implemented at presentation, might affect patient outcomes and the long-term use of resources. This level of information is not available. Hence the scope for formal cost-effectiveness analyses of different primary care treatments or management approaches is limited.

Clinical assessment and evaluation, with identification and modification of risk and precipitating factors, are identified in this guideline as important elements of a baseline assessment in general practice. An exercise electrocardiogram is advocated for its prognostic information, particularly in screening for patients requiring further investigation. The value of all this information in terms of incremental improvement in health outcomes (by subsequent intervention) has not been demonstrated formally.

Patients with stable angina may use a sequence of drugs, beginning with a sublingual glyceryl trinitrate preparation for rapid symptomatic relief. Forms of sublingual glyceryl trinitrate include tablet and aerosol, but since the effect of these may last only 30 minutes, modified-release and percutaneous preparations have been developed. In addition, isosorbide dinitrate and mononitrate are available for prophylactic use. Prophylactic treatment is recommended for patients requiring regular symptomatic relief and should first be attempted with a beta-blocker.

Considerable variations exist in the cost of nitrates, although treatment provided for immediate symptomatic relief cannot necessarily be compared with prophylactic forms. In 1997 in England, the Prescription Pricing Authority reimbursed prescriptions for nitrates in primary care to the value of £60 million (*Table 1*). This amount is totally dwarfed by over £213 million paid out for calcium-channel blockers prescribed for a range of conditions including angina. However, it is of interest to note that two-thirds of the cost of nitrates went in the purchase of isosorbide mononitrate, and nearly all of this accrued because of the use of expensive proprietary modified-release forms. It is unclear how many patients currently receiving isosorbide mononitrate alone could be more appropriately managed on a beta-blocker alone. A simple analysis of volume of use is provided, in terms of script items, although these may not strictly be comparable across classes of drugs.

Beta-blockers themselves differ greatly in price. A recent review of beta-blockers following myocardial infarction found no evidence of improved efficacy or compliance between beta-blockers with different selectivities.

There are considerable long-term costs associated with angina and its sequelae. Appropriate sequencing of drugs, reflecting the evidence of effectiveness presented in this guideline, and avoiding the use of expensive proprietary forms of drugs that provide questionable additional benefits, may ensure that the best use is made of limited resources.

Drug	Reimbursed cost (£ millions)	Script items (thousands)	Cost/item (£)
Nitrates	**60.5**	**8,370**	**7.23**
Glyceryl trinitrate	17.8	3,222	5.51
Isosorbide dinitrate	1.6	529	3.04
Isosorbide mononitrate	41.5	4,616	9.00
Beta-blockers	**78.0**	**14,811**	**5.26**
Calcium-channel blockers	**213.5**	**13,390**	**15.94**

Comment

The guideline development group considered cost issues when specifying appropriate investigations or drug use. The limited evidence available led to the following 'D' level, consensus-based, recommendation concerning drug selection.

- Within any drug class, patients should be treated with the cheapest preparation that they can tolerate, comply with and that controls their symptoms (D).

In the absence of quantified effectiveness data it was not possible to move beyond such statements of (approximate) cost minimisation. (Cost minimisation formally requires alternative treatments to demonstrate equivalence in health outcomes, thus reducing a treatment decision to one of cost alone. In some instances the evidence in the angina guideline is too imprecise to formally show equivalence, and a lower standard of being 'consistent with equivalence' has been applied.) Thus, in the absence of adequate effectiveness data to demonstrate the value of more expensive drugs, the guideline development group was content to allow cost to drive the choice of drug within each drug class. The sequence in which drugs were tried by patients to find adequate symptomatic control was determined by the effectiveness data.

The primary care management of asthma in adults

The aim of this guideline[57] was to provide recommendations to guide primary healthcare professionals in their management of adult patients with asthma. The guideline development group reconvened in 1999 to update the first version of the North of England Asthma Guideline produced in 1996.[58] The areas covered by the guideline were: drug treatment (including devices), exacerbations of asthma, complementary therapies, allergen avoidance, smoking cessation, patient education, self-management, and referral.

Cost-effectiveness and asthma

An analysis of the cost-effectiveness of different strategies for the treatment or care of asthma requires assessment of the relative use of resources (e.g. drugs, contacts in primary and secondary healthcare, time from work) and health consequences (control of asthma, activities of daily living, quality of life). This level of information to inform decisions about care is generally unavailable from trials. There is some uncertainty about the cost-effectiveness of education and management strategies, and whether these should be directed at all patients or at a sub-group of patients. The available evidence and its shortcomings are discussed in the guideline.

With regard to the choice of drug, the guideline takes the position that when drug therapy is indicated, in the absence of reliable evidence to differentiate between different delivery, form, or brand of product, the cheapest forms available should be used.

Utilising data provided by the Prescription Pricing Authority, the most commonly used drugs are shown in *Table 2,* disaggregated by form. It is clear that GPs do not always use the cheapest form of a drug: for example, about 20% (by volume) of β_2-agonists and steroids are prescribed in powder form. When there is no evidence of any differences in therapeutic effect, potentially large savings could be achieved by changing to the cheapest form if patients' circumstances permit.

Comment

As with the angina guideline, there was little scope for formal comparative cost analysis within the guideline. There were four similarly worded recommendations in specific sections of the guideline that addressed economic issues.

TABLE 2 *Drugs prescribed for asthma in primary care (England, 1997)*

Drug and formulation	% of volume prescribed*	Cost (£)/person-year at DDD
Adrenoceptor stimulants (BNF section 3.1.1)		
Salbutamol		
Aerosol MDI	74.6	28
Aerosol other†	14.2	111
Powder	2.1	73
Solution	3.5	236
Tablet	0.5	151
Oral liquid	5.2	44
All forms	*100*	*50*
Salmeterol		
Aerosol MDI	71.8	348
Powder	28.2	377
All forms	*100*	*356*
Terbutaline		
Aerosol MDI	48.1	52
Powder	48.5	116
Solution	1.3	259
Tablet	1.6	50
Oral liquid	0.5	194
All forms	*100*	*89*
Corticosteroids (respiratory) (BNF section 3.2.0)		
Beclomethasone		
Aerosol MDI	79.6	137
Aerosol other†	8.2	130
Powder	12.2	229
All forms	*100*	*148*
Budenoside		
Aerosol MDI	39.6	138
Powder	53.7	270
Solution	6.7	1342
All forms	*100*	*290*
Fluticasone		
Aerosol MDI	68.6	295
Powder	31.4	354
All forms	*100*	*313*

* *Annual data have been estimated from data available for three quarters of 1997*
† *Includes vortex and breath-activated devices*
BNF, British National Formulary (32nd edition);[70] *DDD, defined daily dose; MDI, metered-dose inhaler*

- As there is no good evidence of clinically important differences between different inhaled short-acting β_2-agonists, patients should be treated with the cheapest preparation that they can effectively use (D).
- As there is no good evidence of clinically important differences between different inhaled long-acting β_2-agonists, patients should be treated with the cheapest preparation that they can effectively use (D).
- As there is no good evidence of clinically important differences between differing inhaled corticosteroids, patients should be treated with the cheapest inhaled corticosteroid that they can effectively use and that controls their symptoms (D).
- Healthcare professionals advising patients should prescribe the cheapest drug delivery device that the patient can use and comply with effectively (D).

As in the angina guideline, in the absence of quantified effectiveness data it was not possible to move beyond such statements of (approximate) cost minimisation.

The prevention and treatment of diabetic foot ulcers

The aim of this guideline[59–61] was to provide healthcare professionals with recommendations for the prevention of, and minimisation of the consequences of, foot ulcers in patients with non-insulin dependent diabetes. The scope included: care of the diabetic foot without complications (organisation between primary and secondary care, the role of healthcare professionals, patient education); the foot at raised risk of complications (definition, identification, prevention of complications, patient education); and the ulcerated foot (diagnosis, treatment, patient education).

Cost-effectiveness

Evidence in many areas precluded the formal use of cost data. However in one area formal cost-effectiveness analysis based on a single randomised trial was possible, and in two other areas messages relating to cost-minimisation emerged.

Screening and protection programme for patients at raised risk

McCabe and colleagues[62] report a screening and protection programme conducted in an English outpatient clinic setting which randomised 2001 patients with diabetes. Patients in the intervention group ($n = 1001$) were screened and patients at raised risk ($n = 259$) were recalled. Following a second assessment, 192 (19.2%) patients were entered into a foot protection programme. These patients had gross neuropathy indicated by foot deformities, vascular disease indicated by an ankle-brachial index ≤ 0.75 or a history of ulceration. Patients in the foot protection programme were eligible for weekly clinics that provided chiropody, hygiene maintenance, hosiery and protective shoes as well as education on daily hygiene, clothing and footwear.

Compared with the control group, the intervention group demonstrated non-significant trends in reduced ulceration and minor amputations, and statistically significant reductions in overall and major amputation. Of those patients presenting with ulcers, significantly fewer progressed to amputation in the intervention group ($p = 0.006$), which suggested that ulcers were spotted sooner and treated more effectively.

Setting the costs of intervention against the costs of reduced amputation alone, the authors conclude that the intervention appears to be cost saving. Thus it is possible that the criteria for entrance to the foot protection programme may have been too stringent and a broader inclusion may be acceptably cost-effective.

Monitoring to detect patients at raised risk of ulceration

A range of techniques are available for detecting patients at raised risk of diabetic foot ulcers although only testing the vibration perception threshold using a biothesiometer and sensory testing using a monofilament have been tested prospectively. Kumar and colleagues[63] commented that filaments were easy to use, light (150 g) and cheap (£12/set) when compared with a biothesiometer weighing 2.5 kg, requiring a power source and costing £400. The findings of the available prospective studies and the relative performance in head-to-head studies with surrogate end-points suggest that monofilaments provide a portable and cost-effective alternative in first-line monitoring for neuropathy.

Wound dressing selection for foot ulcers

No robust evidence of the relative effectiveness or cost-effectiveness of any dressing has emerged. The costs of different proprietary dressings, reimbursed by the NHS, are similar: for a 10 cm × 10 cm dressing, alginate dressings cost £1.48–£1.65, foam dressings cost £1.73–£1.98, and hydrocolloid dressings cost £1.01–£2.14. A 10 cm × 10 cm non-proprietary foam dressing costs £1.03 and paraffin gauze costs £0.30. Although it is recognised that the economics of wound dressing are more complex than the cost of the dressings alone, there is no reliable evidence to justify further analysis. The choice between different dressings may therefore depend not only on the type or stage of wound, but also on personal experience, availability of dressing, patient preference or tolerance and the site of the wound.[64]

Comment

Although the guideline development process used qualitative summary, it was possible to make a cost-effectiveness-based recommendation in the one area where there was a large and representative trial. Thus, on the basis of the limited evidence available, the guideline development group reached the following cost-informed recommendations.

- Patients with risk factors of ulceration should be referred to a specialist multidisciplinary protection programme (A).
- Identification of neuropathy based on insensitivity to a 10-g (gauge 5.07) monofilament is convenient and appears cost-effective (D).

- In the absence of strong clinical or cost-effectiveness evidence, healthcare professionals should use wound dressings that best match clinical experience, cost, patient preference and the site of the wound (D).

Chapter 5

Case studies of guidelines using quantitative evidence summary methods and looking at relatively narrow clinical questions

Six guidelines were developed using quantitative summary methods: each set out to address issues of effectiveness, compliance, harm, quality of life, and cost, although useful data were not always available. For each guideline a short content summary precedes a presentation of cost issues and cost-effectiveness. The descriptions of the guidelines conclude with comments on the particular issues raised by the process.

ACE inhibitors in the primary care management of adults with symptomatic heart failure

The aim of this guideline[25,26] was to provide recommendations to guide primary healthcare professionals in their use of ACE inhibitors in adult patients with heart failure. The effectiveness data showed that ACE inhibitors are clinically effective in the treatment of heart failure. In patients with symptomatic heart failure the beneficial effects of ACE inhibitors are demonstrated for those patients with a reported left ventricular ejection fraction of 35% or less; the greater the impairment, the greater the benefit. Long-term treatment trials of patients with prior myocardial infarction and left ventricular dysfunction also indicate a clinically important benefit from ACE inhibition, although there may be risks (hypotension, worsening chest pain) of rapid commencement of therapy (within 3 days of the onset of symptoms). There is an improvement in symptoms and exercise tolerance when patients with symptomatic heart failure and a reported left ventricular ejection fraction of 35% or less are given an ACE inhibitor. The value of the improvements in terms of patients' general well-being is uncertain.

Cost-effectiveness

McMurray and colleagues[65] estimated that the annual direct cost of heart failure to the NHS was £360 million in 1990/1991 (1% of the NHS budget). They commented on the importance of strategies to reduce hospitalisation of patients with heart failure since this accounted for about 60% of total costs.

Trials consistently show a reduction in hospitalisation for progressive heart disease when on ACE inhibitor therapy. In the Studies of Left Ventricular Dysfunction (SOLVD) treatment trial the percentage of patients hospitalised once or more was 36.6% in the placebo group and 25.8% with enalapril.[47] Further analysis indicates that overall heart failure hospitalisation rates fell from 21.9 to 15.4 per 100 patient-years of symptomatic disease (*Table 3*).[66] (This may be an underestimate of the reduction in hospitalisation, since the SOLVD trials only published average treatment duration for both treatment and control groups together. Treatment duration on enalapril is likely to be more – and on placebo, less – than this average.) Since hospitalisation is most frequent in the last stages of disease, it is unclear whether lasting reductions in hospitalisation rates are achieved or to what extent trial findings simply reflect a 'time-window' effect (i.e. more patients on enalapril make it to the end of trial follow-up without heart disease progressing but all patients deteriorate in following years). Another possibility is that while hospitalisations for heart failure decrease, there are hospitalisations for competing causes of morbidity and so the overall hospitalisation rates are not reduced. However, the data do not suggest that greater hospitalisation for other reasons offsets reduced heart failure hospitalisation. In fact, ACE inhibitors seem to have a positive effect on other-cause hospitalisation in patients with symptomatic heart failure (*Table 3*).

It is not generally safe to assume that hospitalisation rates found in trials will be matched in clinical practice. However, the hospitalisation rate in the control arm of the SOLVD treatment trial precisely matches the rate found in general practice in England. Each GP could expect, on average, four inpatient cases with heart failure each year and SOLVD trial data suggest that ACE inhibition might prevent (or delay) one of these hospitalisations.

TABLE 3 *Overall hospitalisation rates from the SOLVD trials*[47,66]

	No. of hospitalisations		Hospitalisation/patient/year*	
Intervention ...	**Enalapril**	**Placebo**	**Enalapril**	**Placebo**
Treatment trial (symptomatic patients)				
HF	683	971	0.154	0.219
All reasons	2396	2833	0.540	0.640
Prevention trial (asymptomatic patients)				
HF	306	454	0.047	0.069
All reasons	2645	2839	0.402	0.430 NS

* *Differences are statistically significant (p < 0.05) unless indicated by NS (not significant)*

HF, heart failure

Several cost-effectiveness analyses have estimated the effect of ACE inhibitors on patient populations[67,68] based on the SOLVD treatment trial. A potential criticism of these models is that they extrapolate and perpetuate the relative improvement achieved by ACE inhibitors beyond the trial window. The most relevant study for the UK setting estimated the costs and benefits of 4 years of treatment with enalapril in primary care.[69] Costs included initialisation, drug treatment, monitoring and hospitalisation, but explicitly excluded investigations such as echocardiography. Survival and hospitalisation rates were based on SOLVD treatment trial data. Findings varied from a cost saving of £11/patient to cost-effectiveness of £2508 per life-year gained. The study considered a number of scenarios reflecting primary care or inpatient initialisation of therapy and different costs of inpatient care. Not all variables of interest were explored and so some uncertainties about the results remain. It is evident that ACE inhibitors reduce hospital costs in the short to medium term (the 41 months average follow-up of the SOLVD treatment trial). A more conservative assumption in the analysis would have been to assume a delay in hospitalisation rather than a lasting reduction.

The annual cost of purchasing ACE inhibitors (at maintenance doses) ranges from £100 to £340 per year (*Table 4*) on the basis of doses reported in the British National Formulary.[70] However, it is unclear whether these maintenance doses are, in all instances, therapeutically equivalent to the trial doses (which are shown in brackets in *Table 4*).

The average cost per patient using ACE inhibitors in primary care may vary from a small cost saving through to a net cost of £1600 over 4 years (see *Table 5*). This represents a fractional overestimate since withdrawal from treatment would lead to lower drug purchase costs. However, for simplicity the guideline development group preferred the presentation shown. Hence it is likely that the cost-effectiveness of ACE inhibitors for heart failure falls in the approximate range £0 to £10,000 per life-year gained, taking into account the assumptions listed and remaining uncertainties. The important variables are the cost of the ACE inhibitor itself and hospitalisation savings. It is not possible in this simple model to explore the influence of compliance with therapy on the cost-effectiveness estimates presented. The trial data, analysed on an intention-to-treat basis, reflect the level of compliance achieved in the SOLVD treatment trial; the degree to which this may be generalised to general practice in the UK is uncertain. Where non-compliance involves ceasing treatment then both costs and benefits are foregone and the cost-effectiveness ratios are not substantially altered. Substantial crossover to ACE inhibitor therapy in the placebo group in the SOLVD trial may mean that the attributable benefits are underestimated.

Comparisons with cost-effectiveness analyses of other health-service interventions should be treated with caution since it is necessary to appraise the methods of these studies before comparing the findings. However, Drummond and colleagues[71] estimated the cost-effectiveness of drug treatment for patients with elevated total serum cholesterol and for whom dietary measures had failed. The cost-effectiveness of simvastatin 20 mg daily, when compared with no intervention, was estimated to be £11,900 to £56,650 per life-year gained, depending on age and pre-treatment cholesterol level. Advocates of the new selective serotonin reuptake inhibitor (SSRI) antidepressants argue for their use on the basis of better safety in overdose. Freemantle and colleagues[72] tentatively explored the routine

TABLE 4 *Annual (maintenance) cost of ACE inhibitors for heart failure*

ACE inhibitor*	Daily dose (mg)†	Brand name (manufacturer)	Pack size (no. of tablets or capsules)	Dose	Pack cost (£)‡	Cost for 30 days (£)	Cost for 1 year (£)
Captopril	50–75	Capoten®	56	25 mg b.d.	12.03	12.89	157
	(50–300)	(Squibb)	56	25 mg t.d.s.	12.03	19.33	235
Enalapril	20	Innovace®	28	20 mg o.d.	13.10	14.04	171
	(2.5–40)	(MSD)	28	10 mg b.d.	11.03	23.64	288
Fosinopril	10–40	Staril®	28	10 mg o.d.	12.04	12.90	157
	(20)	(Squibb)	28	20 mg b.d.	13.00	27.86	339
Lisinopril	5–20	Carace® (Du Pont),	28	5 mg o.d.	9.58	10.26	125
	(20)	Zestril® (Zeneca)	28	20 mg o.d.	13.38	14.34	174
Perindopril	4 (4)	Coversyl® (Servier)	30	4 mg o.d.	13.65	13.65	166
Quinapril	10–20	Accupro®	28	5 mg b.d.	10.30	22.07	269
	(5–40)	(Parke-Davis)	28	10 mg b.d.	10.07	21.58	263
Ramipril	2.5–5	Tritace®	28	2.5 mg o.d.	7.51	8.05	98
	(5–20)	(Hoechst)	28	5 mg o.d.	9.55	10.23	124

* *ACE inhibitors licensed in the UK for treatment of HF and listed in the BNF (32nd edition)*[70]

† *The maintenance dose range cited in the BNF (32nd edition).*[70] *Patients are initiated on lower doses. The range of doses used in trials is shown in brackets*

‡ *As reported in the BNF (32nd edition)*[70]

TABLE 5 *Net cost and benefit per patient of ACE inhibitors for heart failure*

	Optimistic	Conservative
Assumptions (optimistic or conservative)*		
ACE inhibitor £100/year or £340/year for 4 years (see *Table 4*)	£400	£1400
Initiation of therapy by 2 GP visits or 2 outpatient visits†	£20	£138
Reduced hospitalisation or no reduced hospitalisation‡	–£626	£0
GP visits related to HF unchanged or 1 extra visit/year for 4 years§	£0	£40
Net cost range	–£206	£1578
Increased life expectancy (based on placebo comparison)#	0.203 years	0.203 years
Incremental cost-effectiveness of implementing ACE inhibitor therapy¶	Small cost saving and health gain	£7770/life-year gained

* *Costs and benefits that arise from the addition of ACE inhibitors to current care are shown. Diagnosis costs are excluded because of the variation in tests performed, the lack of adequate cost data and because these costs may occur in any case as part of normal care. Costs presented here are overestimates as withdrawal from treatment has not, for simplicity, been included*

† *Cost per GP consultation, £10 (excluding prescribing cost);*[73] *cost per outpatient visit, £69;*[74] *costs of additional blood tests are excluded because no adequate cost data were found*

‡ *Calculation based on: difference in SOLVD trial treatment and control hospitalisation rates (21.9% – 15.4%) x 4 years; inpatient stay of 14.5 days;*[65] *cost per inpatient-day, £166*[73]

§ *Since patients visit their GPs, on average, once a year in relation to HF it is not plausible to assume an optimistic reduction in GP visits although treatment does delay disease progression and associated morbidity*

On the basis of the placebo-controlled findings of the SOLVD treatment trial, improved survival was highly statistically significant (p = 0.0036 by stratified log rank test). However, the survival gain calculation (using Irwin's Restricted Mean) does not provide a useful CI. The point estimate is thus used in optimistic and conservative scenarios.

Future costs and benefits are not discounted because of the short 4-year time frame and because all important costs are distributed along with the benefit in time. Discounting will not substantially alter the cost-effectiveness ratios

¶ *Survival gains are truncated in the SOLVD trial, and it is reasonable to presume that if treatment stopped there would be some additional benefit after cessation of therapy. However this is not modelled since it is probable that therapy would continue and so both costs and benefits would occur after 4 years*

It was not possible to meaningfully explore compliance with therapy (see text)

first-line use of SSRIs in primary care to prevent fatal overdose and estimated cost-effectiveness at £19,000 to £173,000 per life-year gained using a range of assumptions.

Comment

Within this guideline the group were considering the addition of a new and effective drug treatment for a patient group. The new drug had significant acquisition costs. The group's overall conclusion was that "ACE inhibitors appear to be a cost-effective use of resources when compared with other common health-service interventions (III)" and they made the following recommendation.

- Treatment of heart failure with ACE inhibitors is cost-effective (C).

To reach this conclusion, the group were comfortable using simple, explicit modelling of the consequences of treatment with ACE inhibitors. Although the effectiveness data within this modelling were based on trial data, such were the assumptions needed to populate the model that the recommendation was made only at level C. The assumptions were subjected to a 'sensitivity analysis' by examining both optimistic and conservative scenarios, thereby allowing the group to see the likely limits of treatment cost-effectiveness estimates.

Aspirin for the secondary prophylaxis of vascular disease in primary care

The aim of this guideline[27,28] was to provide recommendations for GPs on the use of aspirin for the secondary prophylaxis of non-fatal and fatal cardiovascular disease and stroke in the management of adult patients at raised vascular risk. The summary of the effectiveness data suggested the following.

- Aspirin initiated within 24 hours of acute myocardial infarction lowers the risk of a vascular event over the subsequent month.
- Aspirin lowers the risk of a subsequent vascular event if given to patients who have: previously had a myocardial infarction; stable angina; unstable angina; a past history of transient ischaemic attack or mild to moderate stroke.
- Aspirin given to patients with intermittent claudication or diabetes appears to have a small and statistically uncertain effect upon the risk of experiencing a subsequent vascular event.

Having defined the efficacy of aspirin and addressed issues of dose and duration of treatment, the group considered issues of cost.

Side-effects and costs of aspirin

Trials of aspirin to prevent cardiovascular disease address clinical end-points and offer no direct evidence about the impact of aspirin on patients' quality of life, although aspirin has two potentially serious side-effects: intra-cerebral haemorrhage and gastrointestinal bleeding. The value of aspirin to individual patients must balance the increased risk of haemorrhage against the reduced likelihood of a cardiovascular event. Gastrointestinal problems occur for reasons besides use of aspirin, and placebo-controlled trials permit the calculation of the risk attributable to aspirin. A recent review examined all trials listed in the Antiplatelet Trialists Collaboration[75] for information on toxicity.[76] Patients with a history of peptic ulcer, gastrointestinal bleeding or contraindication to aspirin were generally excluded from trials. Twenty-one trials with 20,011 patients randomised to aspirin gave a total of 76,215 years of exposure. Comparing patients receiving aspirin with those receiving placebo, the pooled odds ratio for all forms of gastrointestinal bleeding was 2.0 (95% CI, 1.5 to 2.8) and for bleeding leading to hospitalisation it was 1.9 (95% CI, 1.1 to 3.1). Similarly, when these groups were compared for either peptic ulcers or gastrointestinal symptoms leading to treatment withdrawal, the pooled odds ratios were 1.3 (95% CI, 1.07 to 1.6) and 1.5 (95% CI, 1.1 to 1.9), respectively.

The review found a consistent tendency of lower rates of adverse events at lower doses. In the Aspirin Myocardial Infarction Study,[77,78] which used a daily dose of 1000 mg aspirin, the odds ratio for hospital admission for peptic ulcer reported for patients with previous myocardial infarction was 4.1. In the UK Transient Ischaemic Attack trial[79,80] (patients with previous transient ischaemic attack) the attributable rate of gastrointestinal bleeding was 2.5 and 7.7 per 1000 person-years of treatment with 300 mg and 1200 mg aspirin daily, respectively.

There are two major trials that have used a dose of 75 mg aspirin daily; a summary of reported side-effects is shown in *Table 6*. Thus, it is possible to project the major benefits and risks attributable to 75 mg aspirin daily, but the calculations presented should be treated with caution, since the trials were not adequately powered to measure adverse effects at conventional levels of statistical significance.

TABLE 6 *Events attributable to aspirin (derived from major trials using 75-mg daily dose)*

Trial	Indication	Mean treatment duration	Number enrolled	Events (no. per 1000 patient-years of treatment)* End-point or event	Aspirin	Placebo	Attributable
SALT[81]	Previous transient ischaemic attack or minor stroke	Aspirin: 30.6 months Placebo: 27.5 months	Aspirin: 676 Placebo: 684	Primary end-point†	80.1	109.1	–29.0
				Adverse events:			
				• Total‡	85.3	78.5	6.8
				• Gastrointestinal (excluding bleeding)			
				– Non-severe	37.1	35.1	2.0
				– Severe	12.2	11.5	0.7
				• Haemorrhagic			
				– Non-severe	16.8	8.3	8.5
				– Severe	11.6	5.7	5.9
				• Other			
				– Non-severe	12.8	19.8	–7.0
				– Severe	5.2	7.0	–1.8
				Withdrawal due to:			
				• Adverse experience	9.9	6.4	3.5
				• Any reason (excluding primary end-points)	66.7	100.2	–33.5
SAPAT[82]	Stable angina	Aspirin: 49.9 months Placebo: 50.2 months	Aspirin: 1009 Placebo: 1026	Primary end-point§	19.3	28.9	–9.6
				Adverse events:			
				• Haemorrhagic			
				– Minor	1.7	0.7	1.0
				– Major	2.6	1.9	0.8
				– Fatal	2.1	1.2	1.0
				Withdrawal due to:			
				• Adverse event	26.0	23.3	2.7
				• Any reason	86.3	97.2	–10.9

* *Numbers shown are rounded to one decimal place*
† *Stroke (fatal or non-fatal) or non-stroke death*
‡ *Some patients had more than one type*
§ *Myocardial infarction (MI; fatal or non-fatal) or sudden death*

SALT, Swedish Aspirin Low-dose trial; SAPAT, Swedish Angina Pectoris Aspirin Trial

Assuming 1000 person-years of treatment, the following effects are attributable to aspirin 75 mg daily.

- Ten patients with stable angina will avoid vascular events (non-fatal or fatal myocardial infarction or sudden death). However, one patient will suffer a fatal bleed, one patient will suffer a major non-fatal bleed and one patient will experience a minor bleed (where 'bleed' includes stroke and gastrointestinal haemorrhage).
- Twenty-nine patients with a previous transient ischaemic attack or minor stroke will avoid a vascular event (non-fatal or fatal stroke or other vascular death). However, six patients will suffer a serious bleed (possibly fatal) and eight patients will suffer less serious bleeds. The higher rate of bleeds in the SALT trial[81] reflects, in part, a greater occurrence of intracranial bleeds; these were comparatively uncommon in the SAPAT[82] trial.

Presumably because aspirin is so cheap to purchase, we found no adequate economic analyses evaluating its use. The cost of aspirin itself is negligible. Non-proprietary aspirin, in 75-mg dispersible tablets, costs approximately £1 per year to prescribe, although proprietary brands may cost 10–20 times more. A formal analysis would weigh up the overall benefit of the decision to prescribe aspirin against its net cost. From a health-service perspective, the net cost includes the cost of aspirin, treatment for attributable adverse events and savings from fewer vascular events. Patients with vascular disease tend to consult regularly with their GPs and it is likely that any increase in GP consultation due to treatment with aspirin would be small. Since the reduction in vascular events considerably exceeds the attributable adverse events, and given the nature of the medical interventions for both, it is likely that aspirin treatment results in a net cost saving to the health service. The balance of costs could shift adversely if it was necessary to provide expensive antisecretory drugs to ameliorate gastrointestinal symptoms in a significant proportion of patients, although the rates of attributable adverse events reported do not indicate that aspirin 75 mg daily causes a substantial rise in the need for anti-secretory drugs. Cessation of therapy is most likely in the presence of adverse symptomatology. It is likely that the use of aspirin is cost saving or cost neutral (i.e. not involving an increase in healthcare costs in total) although formal cost calculation has not proved possible because there are inadequate hospital cost data.

Comment

In this guideline the group were dealing with the addition to current care of a drug that was both effective and had a low acquisition cost. The group summarised the data in the following manner.

- The benefits of prophylactic use of aspirin in the secondary prophylaxis of vascular disease considerably outweigh the attributable risks of a gastrointestinal or cerebrovascular bleed; the use of aspirin is likely to be cost saving or cost neutral (I).

This led to a recommendation to use aspirin.

First-line drug treatment for depression in primary care

The aim of this guideline[22,83] was to provide recommendations to guide primary healthcare professionals in their use of antidepressants in the treatment of adults with depression and for whom the agreed course of action is to prescribe. The effectiveness data showed that tricyclic antidepressants were slightly more efficacious than SSRIs or related drugs, although this effect was of uncertain practical importance. SSRIs and related drugs were slightly better tolerated than tricyclic antidepressants, reducing the risk of drop-out by about 4% during 6 weeks of treatment in double-blind randomised trials. There was a substantial range of toxicity associated with different antidepressants currently used in primary care. The SSRIs and lofepramine were associated with the smallest risk of fatal poisoning.

Economics of antidepressants

Data on reimbursements in England for all antidepressants show that SSRIs cost five to six times more to purchase than tricyclics. However, higher acquisition costs for newer pharmaceuticals can be justified if they are offset by reduced costs in other parts of the health system or produce additional health gains (which may themselves generate further productivity benefits in the economy). This has been the argument put in favour of the SSRIs: cost savings due to reduced hospitalisations roughly compensate for their increased acquisition cost and so there is no reason to discriminate against them in the first-line treatment of depression on the basis of cost. Reduced use of health-service resources could occur through several mechanisms. The small (but statistically significant) difference in drop-out rates found in trials could reflect a better tolerability, leading to fewer patients with

treatment failure presenting as outpatients or as inpatient admissions. Greater safety in overdose may lead to fewer hospital admissions for poisoning-related incidents. Also, careful matching of pharmacological properties (e.g. level of sedation) to patient's lifestyle may reduce drug-associated accidents.

It is possible by looking at current admission statistics to estimate the total cost of admissions for poisonings and neurotic disorders and so set upper and lower bounds on any potential savings achievable by antidepressant choice.

In England, in 1994–1995, there were 33,048 ordinary admissions for neurotic and personality disorders with a mean length of stay of 30 days (indicating 991,076 bed-days). Since this disease category is very broad, these numbers are used as an upper bound for inpatient care for depression in England each year. Day-case admissions are reported to be negligible. There were 37,460 psychiatric outpatient episodes in Scotland in 1991 and each new outpatient averaged 7.2 attendances. No comparable data have been located in the public domain for England, but assuming the Scottish data are generalisable, they indicate 355,000 treatment episodes per year in England.

Two scenarios (*Boxes 13* and *14*) are presented to explore the likely impact of general policies

BOX 13 'Conservative scenario' of the likely impact of increased use of SSRIs or lofepramine and decreased use of other tricyclic antidepressants as first-line treatment for depression in primary care

Assumptions

- Accidental fatal poisonings associated with a single-ingested tricyclic antidepressant will reduce proportionately as tricyclic use is reduced (using low estimate of poisonings)
- All hospitalisations attributed to the toxic effects of antidepressants will reduce proportionately as tricyclic use is reduced (using low estimate of hospitalisations)
- Differences in efficacy and drop-out between antidepressants are insignificant (or approximately cancel out one another in consequences to patients) leading to no net change in primary care, outpatient use or inpatient psychiatric services as a result of antidepressant choice

For every patient-year of treatment changed

A	Cost of SSRI*	Average cost of SSRI/year	(£)	282
B	Cost of lofepramine*	Average cost of lofepramine/year	(£)	101
C	Cost of tricyclic*†	Average cost of tricyclic/year	(£)	45
D	Cost of toxicity admission‡	35,140 bed-days × £160 per day/528,700	(£)	11
E	A – (C + D)	SSRI net cost/patient	(Δ£)	226
F	B – (C + D)	Lofepramine net cost/patient	(Δ£)	45
G	Lives saved (SSRI)§	0.000090 – 0	(ΔLS)	0.000090
H	Lives saved (lofepramine)	0.000090 – 0.000003	(ΔLS)	0.000087
	E/G	Incremental cost per life saved by switching to an SSRI from an older tricyclic	(Δ£/ΔLS)	2,500,000
	F/H	Incremental cost per life saved by switching to lofepramine from an older tricyclic	(Δ£/ΔLS)	520,000
	(E – F)/(G – H)	Incremental cost per life saved by switching to an SSRI from lofepramine	(Δ£/ΔLS)	60,000,000

* *At WHO DDD*

† *Average yearly cost of a tricyclic or related antidepressant* ***excluding*** *lofepramine*

‡ *Based on average cost per inpatient-week in Scotland 1995/1996*[84] *and converted to a daily rate*

§ *Differences in fatality association rates between SSRIs and lofepramine are not statistically significantly different*

BOX 14 'Optimistic scenario' of the likely impact of increased use of SSRIs or lofepramine and decreased use of other tricyclic antidepressants as first-line treatment for depression in primary care

Assumptions

- All fatal poisonings associated with tricyclic antidepressants are preventable and will reduce proportionately as tricyclic use is reduced (high estimate)
- Hospitalisations for toxic effects, accidents and falls attributed to antidepressants will reduce proportionately as tricyclic use is reduced (high estimate)
- SSRIs are as efficacious as the tricyclics but result in 4% fewer drop-outs: 5% of drop-outs are assumed, regardless of drug, to be admitted to hospital for an average of 30 days, and 5% are assumed to make use of an average 7.2 outpatient visits

For every patient-year of treatment changed

A	Cost of SSRI[*]	Average cost of SSRI/year	(£)	282
B	Cost of lofepramine[*]	Average cost of lofepramine/year	(£)	101
C	Cost of tricyclic[*†]	Average cost of tricyclic/year	(£)	45
D	Cost of poisoning admissions[‡]	35,140 bed-days × £160 per day/528,700	(£)	11
E	Cost of admissions due to accidents[‡]	84,400 bed-days × £160 per day/528,700	(£)	26
F	Cost of psychiatric admissions[§#]	4% × 5% × 30 × £100/day	(£)	6
G	Cost of outpatient attendences[§¶]	4% × 5% × 7.2 × £40/visit	(£)	1
H	A – (C + D + E + F + G)	SSRI net cost/patient	(Δ£)	194
I	B – (C + D)	Lofepramine net cost/patient	(Δ£)	45
J	Lives saved (SSRI)[**]	0.000808 – 0.000041	(ΔLS)	0.000767
K	Lives saved (lofepramine)[**]	0.000808 – 0.000062	(ΔLS)	0.000746
	H/J	Incremental cost per life saved by switching to an SSRI from a tricyclic	(Δ£/ΔLS)	250,000
	I/K	Incremental cost per life saved by switching to lofepramine from a tricyclic	(Δ£/ΔLS)	60,000
	(H – I)/(J – K)	Incremental cost per life saved by switching to an SSRI from lofepramine	(Δ£/ΔLS)	7,100,000

[*] *At WHO DDD*

[†] *Average yearly cost of a tricyclic or related antidepressant* ***excluding*** *lofepramine*

[‡] *Based on average cost per inpatient-day in Scotland 1995/1996*[84]

[§] *The cost of treating one drop-out from either a tricyclic or an SSRI is assumed to cost, on average, the same. Then it is only the costs of additional drop-outs for those on a tricyclic over and above those dropping out from an SSRI that present additional costs. The additional drop-out of 4% comes from the meta-analysis of drop-out. The estimates of 5% of drop-outs becoming inpatients and outpatients is based on expert panel data in the literature.*[85] *These estimates can be tested assuming that 30% of the 3.7 million antidepressant treatment episodes per year in England (see* Table 3*) result in drop-out: 56,000 inpatient and outpatient treatment episodes would be predicted per year in England. Since only 33,048 ordinary admissions were recorded for neurotic and personality disorders in 1994/1995, inpatient admission may be overestimated. 56,000 outpatient episodes represents about one-sixth of all yearly psychiatric outpatient episodes*

[#] *Based on the average cost per inpatient-week for mental illness in Scotland in 1995/1996*[84] *and converted to a daily rate*

[¶] *Based on the average cost per outpatient attendance for mental illness in Scotland in 1995/1996*[84]

[**] *Differences in fatality rates between SSRIs and lofepramine are not statistically significantly different*

to increase the use of SSRIs or lofepramine and decrease the use of other tricyclic antidepressants upon NHS costs and toxicity-associated fatalities.

It has been claimed that choosing SSRIs for the treatment of depression in primary care will not increase overall healthcare costs: two scenarios have explored this claim. There is currently insufficient evidence to conclude that changing the choice of antidepressant in UK primary care may substantially affect toxicity-associated or accident-related hospitalisations or cause a reduction in fatalities. Thus, the scenarios presented are tentative, exploring different beliefs by extrapolating and interpolating available data. The guidelines group thought that the assumptions used in the two scenarios explored the reasonable bounds of plausibility, although if increased SSRI use is demonstrated to lead not just to reduced accident admissions but also to reduced accident fatalities, the case for the cost-effectiveness of SSRIs should be reviewed. It was thought inappropriate to explore this possibility, since some tricyclics are only slightly sedating whereas patients receiving SSRIs sometimes require adjunctive sedative therapy. Benefits, if achievable, should largely be obtained by the appropriate choice of tricyclic antidepressant.

The weakness of the modelling approach of cost-effectiveness analysis can be seen in one published analysis that claimed to show that paroxetine was more cost-effective than imipramine[85] and which caused considerable debate.[86,87] The work of other analysts who revisited the key assumptions of the model led to opposite conclusions.[88] One large pragmatic trial conducted in the USA has attempted to resolve the issue of overall healthcare costs of using different antidepressants,[89] but because of design limitations, its findings have no obvious interpretation in the UK setting.[46,54]

Ideally, life-years gained (a common metric in economic evaluations) could be gauged from estimates of lives saved. The average age of death due to antidepressant fatal poisoning is 39 years for men and 46 years for women, with a population average remaining life-expectancy of about 35 years for both.[90] However, this is likely to be an overestimate of life-expectancy for this patient group because of co-morbidity and the remaining risk of future toxic overdose. To measure accurately life-years gained would require a lifetime disease and intervention model, for which there are no adequate data.

Although fatal poisoning with tricyclic antidepressants is a rare event, hospitalisation attributable to the unwanted effects of these drugs may lie between 3% and 5% per year of treatment. As a strategy for saving life, a general policy of switching from tricyclics to SSRIs does not appear to be cost-effective. When there is a concern about toxicity, benefits may be achieved in a more cost-effective manner through switching to lofepramine.

Comment

The guideline development group was comparing drugs from two different groups that were of similar effectiveness and tolerability but that had considerably different toxicity profiles and drug acquisition costs. The question arising from the guideline process was 'how much should be paid to avoid tricyclic associated poisoning deaths?' The group concluded that a general policy of switching from tricyclics to SSRIs did not appear to be cost-effective and when the toxic effects of tricyclic antidepressants gave cause for concern, replacement by lofepramine appeared to be relatively cost-effective. As a consequence the group made the following recommendation.

- As they represent the most cost-effective option, tricyclic antidepressants should be used as the routine first-line drug treatment for depression in primary care (C).

However, aware of the realities of daily general practice, the group went on to recommend the following.

- The choice of antidepressant should be based on individual patient factors; these would include (D):
  - the desirability or otherwise of sedation or other effects associated with a particular drug
  - previous response to a particular drug
  - co-morbid psychiatric or medical conditions
  - concurrent drug therapy.
- If the toxic effects of the older tricyclic antidepressants are perceived to be a problem, for example in a patient who has previously taken a drug overdose, then lofepramine is a more cost-effective choice than an SSRI (C).

Non-steroidal anti-inflammatory drugs (NSAIDs) versus basic analgesia in the treatment of pain believed to be due to degenerative arthritis

The aim of this guideline[23,24] was to provide recommendations to guide primary healthcare

professionals in the appropriate use of NSAIDs for joint pain believed to be due to degenerative arthritis. The evidence on efficacy and side-effects was summarised as follows.

- In randomised comparative trials, NSAIDs have been shown to reduce pain in patients with osteoarthritis or similar joint pain, compared with simple analgesia alone, although the relative benefits are not large. Many patients in the trials have satisfactory pain relief from simple analgesia.
- There are risks of upper gastrointestinal side-effects associated with NSAID use. The use of NSAIDs in the treatment of pain in primary care should take into account the trade-off of risks and potential benefits, and available alternative analgesics.
- Paracetamol and codeine combined appear to have a slightly greater analgesic effect than paracetamol alone. A combination of paracetamol and dextropropoxyphene also demonstrates small and uncertain benefits over paracetamol alone. However, both combinations are associated with a substantial increase in side-effects.
- H_2 blockers, misoprostol and proton pump inhibitors reduce the risk of NSAID-induced duodenal ulcers. Misoprostol also reduces the risk of other serious upper gastrointestinal injury. H_2 blockers may have a small impact upon severe gastric symptoms in patients taking NSAIDs, though it is not clear that benefits generally exceed those from antacids. Omeprazole appears more effective than misoprostol in reducing abdominal pain in patients treated for NSAID-induced ulcer.

Economic analysis

It is important to consider the net cost of a treatment decision alongside its health benefits, rather than just the purchase cost of treatment (the reimbursed price of the NSAIDs). This may include changes in the use of health-service resources (e.g. hospitalisation) and broader social consequences (e.g. time off work or routine activities).

Paracetamol remains a cost-effective alternative to any NSAID: it has a lower purchase cost and relative absence of gastrointestinal toxicity, while displaying similar levels of patient withdrawal from treatment. Nonetheless, concern remains about the unquantified risk of hepatic damage in overdose with paracetamol.

For the three NSAIDs for which data are available from randomised controlled trials with comparison against simple analgesia, a meta-analysis of non-randomised studies suggests an ordering, on safety grounds, of ibuprofen, diclofenac and then naproxen.[91] However, evidence of the rate of gastrointestinal injury is inadequate to discriminate between diclofenac and naproxen. Whereas diclofenac and naproxen are similarly priced, ibuprofen is three to four times cheaper to prescribe given the forms in which these drugs are currently dispensed. Therefore, in the likely event that ibuprofen results in lower gastrointestinal injury and symptomatology, and without clear evidence of a general therapeutic advantage for naproxen or diclofenac, ibuprofen is the most cost-effective first-line NSAID.

The purchase costs of different preparations of the same NSAID available on the NHS vary widely. There is no good evidence to support the use of more expensive preparations over cheaper ones or the use of the modified-release preparations.

In patients requiring NSAID treatment, it is important to consider what strategies may be available to minimise the risk of gastrointestinal injury. Such preventative strategies should not be confused with treatment of (common) dyspepsia for which prescription or over-the-counter purchase of antacids may be considered when NSAID treatment cannot be modified.

Modelling the cost-effectiveness of misoprostol prophylaxis

A systematic review of economic analyses of the prophylactic use of misoprostol has recently been published.[92] Four of the analyses retrieved concerned use of misoprostol by patients with osteoarthritis[93–96] and one concerned use by patients with rheumatoid arthritis.[97] A cross-national comparison study was also retrieved.[98] All analyses reviewed were found to be based on a trial of 420 American osteoarthritis patients with NSAID-associated abdominal pain who were randomised to misoprostol 100 µg, 200 µg or placebo four times daily for 3 months.[99] Gastric ulcers occurred less frequently ($p < 0.001$) with misoprostol (endoscopically detected lesions > 0.3 cm diameter – placebo, 21.7%; misoprostol 100 µg, 5.6%; misoprostol 200 µg, 1.4%; lesions > 0.5 cm diameter – placebo, 12.3%; misoprostol 100 µg, 4.2%; misoprostol 200 µg, 0.7%). Approximately equal proportions of patients were taking ibuprofen, piroxicam or naproxen. There were no significant differences between groups in duodenal ulceration or adverse effects (diarrhoea, dyspepsia, flatulence, abdominal pain and nausea) except for diarrhoea (placebo,

13.0%; misoprostol 100 µg, 25.1%; misoprostol 200 µg, 39.2%).

Endoscopically detected ulcer rates found in trials may bear little relation to the much lower rates of symptomatic presentation in clinical practice. Clearly, any intermediate outcome has an uncertain relationship with health gain, as valued by quality and quantity of life; this may be more so in the artificial confines of a trial. None of the patients in the trial reported by Graham and colleagues[99] suffered a major complication, haemorrhage or perforation. Each economic analysis made different assumptions about the absolute reduction in symptomatic ulcers and thus conflicting results were obtained, although even the studies with pessimistic assumptions featured rates of reduction in ulcers that were six to seven times that found in a large American rheumatoid arthritis trial.[100]

The cost of misoprostol prophylaxis at the lower dose recommended in the British National Formulary[64] (400 µg/day) is about £135–£140 per patient per year, whether prescribed as additive therapy or in combination with diclofenac (comparing non-proprietary diclofenac sodium 50 mg with Arthrotec 50®). The annual purchase cost of the higher dose used in trials (800 µg/day) is £270–£280 per patient.

A rate of serious gastrointestinal events necessitating hospitalisation for rheumatoid arthritis patients on NSAID therapy can be derived from the control group, receiving placebo, in a recent large US study.[100] The reduction in hospitalisation is estimated from the treatment group receiving misoprostol (*Table 7*).

Hence, for 1000 patient-years of treatment, 7.6 events will be prevented (95% CI, 0.4 to 15.1), at a purchase cost of misoprostol of £230,000 (using the average daily dose of 680 µg reported in the trial). The cost/event prevented is calculated using the CI of events prevented to provide low and high estimates (*Table 8*).

In the trial reported by Silverstein and colleagues,[100] the rate of serious gastrointestinal complications in the control group was 1.9% per person-year of treatment. Extrapolation to the number of person-years of treatment currently prescribed in England indicates 30,000 hospitalisations for NSAID-associated gastrointestinal injury per year. Thus half of the 60,000 annual hospitalisations associated with gastrointestinal ulcer/bleeding in England (Hospital Episode Statistics, 1994–1995) can be estimated to be NSAID-associated. There were 4304 gastrointestinal ulcer-associated deaths (International Classification of Diseases (ICD) 531-3) in England in 1991. Assuming that chance of fatality following hospitalisation is independent of the underlying reason for gastrointestinal injury, then 2150 deaths per year can be attributed to NSAID-associated injury, or 1.38 deaths per

TABLE 7 *Rates of serious gastrointestinal events with and without misoprostol in rheumatoid arthritis patients taking NSAIDs (derived from Silverstein,* et al., *1995[100])*

Treatment	No. of patients	Duration of follow-up	Person-years on drug	Events*	Event rate per year
Misoprostol†	4404	6 months	2202	25	0.0114
Placebo	4439	6 months	2220	42	0.0189

* *Serious gastrointestinal events definitely attributable to NSAID use*
† *Average dose 680 µg*

TABLE 8 *Net cost and serious gastrointestinal events prevented with misoprostol prophylaxis (for 1000 patients)*

Scenario	Events avoided	Cost of misoprostol (£)	Savings (£) from reduced hospitalisation*	Net cost (£)	Cost/event avoided (£)
High	0.4	230,000	1,200	228,800	572,000
Best guess	7.6	230,000	22,800	207,200	27,300
Low	15.1	230,000	45,300	184,700	12,200

* *The average cost of inpatient hospitalisation across all specialities was £3000/episode for Scotland in 1995/1996[84]*

year in 1000 patients taking NSAIDs. This suggests that nearly one in ten serious gastrointestinal complications is fatal. The use of misoprostol led to a 40% relative reduction in serious events (95% CI, 64% to 2%) and so may be assumed to lead to a 40% reduction in the average fatality rate. These figures are used to estimate the cost-effectiveness of a general policy of misoprostol prophylaxis in terms of the cost per life saved (*Table 9*). High and low estimates are derived assuming lives saved are a constant fraction of serious events avoided.

TABLE 9 *Modelled cost-effectiveness of misoprostol prophylaxis (for 1000 treated patients)*

Scenario	Net cost (£)	Lives saved*	Cost/life saved (£)
High	228,800	0.0276	8,290,000
Best guess	207,200	0.552	375,000
Low	184,700	0.883	209,000

* *High, best guess and low estimates calculated as 2%, 40% and 64%, respectively, of 1.38*

The above estimates must be viewed as tentative given the assumptions required to reach them. For example, the relationship between hospitalisation and mortality for ulcers of different underlying cause is unknown; there is currently no direct evidence of gastrointestinal injury-associated death being prevented by misoprostol prophylaxis.

The mean ages of death in men and women due to ulcer, haemorrhage and perforation (ICD-9: 531-3) are 76 and 81 years, respectively. The average life-expectancy in the normal population for both sexes at these ages is about 8 years.[90] Hence, a crude calculation of cost per year of life gained is potentially possible using estimates in *Table 9*. The age distribution of ulcer fatalities presented in national statistics aggregates those ulcers caused by NSAIDs, those related to *Helicobacter pylori*, and those due to other causes. These would need to be disaggregated and the assumptions validated before formal calculation is possible.

On the available evidence, it is not demonstrated that a strategy of routine and unselected misoprostol prophylaxis for patients taking NSAID therapy is cost-effective. Patient review and sequential therapy selection, beginning with simple analgesia, is likely to minimise adverse event rates in the general patient group.

It is possible, although not demonstrated, that misoprostol prophylaxis may be more cost-effective in a high-risk group for which current NSAID therapy has to be maintained. The study by Silverstein and colleagues[100] of patients with rheumatoid arthritis suggested greater relative risks of serious gastrointestinal injury for patients with age > 75 years (odds ratio = 2.48), and for patients with history of peptic ulcer (odds ratio = 2.29), gastrointestinal bleeding (odds ratio = 2.56) or heart disease (odds ratio = 1.84). These risks factors have been presented in such a manner that it is not possible to calculate absolute reductions in the rates of serious events for each high-risk group, and the numbers of events in each case are small. Without absolute risk reductions, neither can we calculate cost per life saved for high-risk groups. However, none of the risk factors appears very important, and the cost-effectiveness of misoprostol prophylaxis in high-risk NSAID-user groups remains undemonstrated.

Although it appears likely that omeprazole may be similar in effectiveness to misoprostol in NSAID-induced ulcer prophylaxis and healing, and also possibly better tolerated (although purchase costs are also higher), trials that rely on detecting ulcers by endoscopy overestimate the effectiveness of protective agents in practice. Also, no large pragmatically designed trials with serious gastrointestinal events as primary outcome are available for omeprazole. Without such data it is not possible to recommend the routine use of omeprazole prophylaxis as an evidence-based strategy.

Comment

In this guideline the group were choosing between simple analgesia and the use of a drug group that had little therapeutic advantage but considerably more side-effects. To understand the evidence it was necessary to perform (relatively simple) modelling of the consequences around the commonest serious side-effect, gastrointestinal haemorrhage. As a consequence of this the group made the following recommendations.

- In terms of cost-effectiveness, patients presenting with painful joints believed to be due to degenerative arthritis should initially be treated with paracetamol. If inadequate symptomatic relief is obtained then ibuprofen is the most cost-effective alternative (C).
- Modified-release NSAID preparations are relatively expensive while no evidence demonstrates that they are more effective than standard therapy; therefore they should not be used (D).

- Prophylactic gastrointestinal protective therapy (with misoprostol or proton pump inhibitors) should not be used routinely, as it is not cost-effective for the reduction of serious gastric events (D). There are a group of patients who are at higher risk of upper gastrointestinal bleeding or perforation for whom prophylaxis may be cost-effective but further evidence is required (D).

The primary care management of dementia

A guideline addressing the management of patients with dementia[29,30] examined the role of donepezil hydrochloride. Donepezil is a piperidine-based derivative that is chemically distinct from other cholinesterase inhibitors, and was developed specifically for the treatment of Alzheimer's disease. An important aim in the pharmacological development of such a drug would be to achieve a therapeutic level of cholinesterase inhibition but without the toxicity and side-effects experienced with previous drugs. The available trials assessed a range of cognitive function tests, but only one study[101] assessed performance using an 'activities of daily living' scale.

The evidence of effectiveness was summarised in the following statements.

- Donepezil has demonstrated a moderate effect upon cognitive function in short-term treatment trials of patients with mild to moderate Alzheimer's disease (I).
- These changes in cognitive function have not been accompanied by measured changes in quality of life, and there is inadequate evidence on the effects on activities of daily living (I).
- Although side-effects increase with dosage, there is evidence that a 5-mg dose is similar in efficacy to a higher dose, and not associated with substantial side-effects likely to lead to withdrawal from therapy. There is no evidence of hepatotoxicity at the doses used in the trials (I).

Economic impact of prescribing donepezil

Alongside the systematic appraisal of effectiveness, compliance, and safety, it is important that implications for health-service resource use and costs should be explored. If possible, broader social costs should also be explored to reflect the carer-dependent nature of the disease. However, the long-term health and resource consequences of drug therapy for Alzheimer's disease are unknown and so the following presentation is partial and speculative, drawing on health-service data.

Alzheimer's disease and health-service resources

Precise estimates of the prevalence of Alzheimer's disease are not directly available at the national level. The two most relevant and widely used disease codes are senile and pre-senile organic psychotic conditions (ICD-9: 290), which mainly feature senile dementia, simple type (ICD-9: 290.0), and other cerebral degenerations (ICD-9: 331), which consists mainly of Alzheimer's disease (ICD-9: 331.0). The relevant hospital activity diagnostic codes are H210, senile and pre-senile organic psychotic conditions (which maps precisely to ICD-9: 290) and H222, other degenerative and hereditary disorders of the central nervous system (mapping to ICD-9: 330, 331, 333–336).

Using the 1991–1992 Morbidity Statistics in General Practice,[102] it can be estimated that in England every year 92,000 patients consult a GP about senile and pre-senile organic psychotic conditions, making about 190,000 consultations. The report briefly reports the number of patients with senile dementia (ICD-9: 290.0), of which there are an estimated 60,000. This suggests that statistics reported for the broader group could be reduced to 60/92 (= 65%) as a more accurate estimate of resource use attributable to Alzheimer's disease. About 23,000 patients made 49,000 consultations coded as other cerebral degenerations (ICD-9: 331). As a cause of death about 80% of events in this patient group are due to Alzheimer's disease (ICD-9: 331.0), and this is used to reduce resource use attributed to the broad group. The estimated caseload for GPs in England is shown in *Table 10*.

Hospital Episode Statistics indicate that for senile and pre-senile organic psychotic conditions there were 46,249 ordinary admissions requiring 3,032,230 bed-days in England in 1993–1994. These figures are reduced to 65% to disregard bed-days for patients with simple dementia. Similarly, for other degenerative and hereditary disorders of the central nervous system, there were 15,396 ordinary admissions requiring 385,438 bed-days. However, this patient group covers a broader range of ICD codes, of which as a cause of death Alzheimer's disease represents 49%: this is used as a correction factor for resource use. Day cases were negligible for both groups. No accurate data are available for outpatient attendance, although 27,590 patients recorded 1,439,425 attendances in England in 1995–1996 for old-age psychiatry.

TABLE 10 *Estimated annual use of NHS resources associated with Alzheimer's disease (England)*

	ICD-9	No. of patients	No. of GP consultations	No. of admissions	No. of bed-days
Senile dementia	290.0	60,000	124,000	30,000	1,971,000
Alzheimer's disease	331.0	18,000	39,000	8,000	189,000
Total	–	78,000	163,000	38,000	2,160,000

The OHE Compendium of Health Statistics[103] lists the cost of an inpatient-day and an outpatient attendance as £123 and £69, respectively, in England, in 1992–1993. A GP consultation is estimated to cost £12.[104]

An estimate of the cost of NHS resources used to treat patients with a diagnosis of probable Alzheimer's disease may be obtained by weighting the resource by the above costs. Including ordinary hospital admissions and GP consultations, the annual estimated NHS cost for England is £268 million in 1992/1993 prices (163,000 × £12 + 2,160,000 × £123). It is unknown what proportion of outpatient attendances for old-age psychiatry may be required for Alzheimer's disease. Arbitrarily assigning 10% of outpatient attendances would raise the total cost to £278 million. The average cost to the NHS per patient per year is thus estimated to be £3560 (in 1992/1993 prices). It is uncertain whether in its use of resources Alzheimer's disease is similar to or more or less intensive than other conditions with which it is grouped. The cost of Alzheimer's disease to the NHS in England is tentative given the list of assumptions involved in its calculation.

A burden-of-illness study estimated the total cost of care (including health and social services) of Alzheimer's disease to be £1039 million in the United Kingdom in 1990/1991, of which £266 million was for direct healthcare costs.[105] The biggest cost component was for residential care, which accounted for 66% of the total. The costs of residential care in private, voluntary, and local authority residential homes have approximately doubled (1996/1997 costs) since this study.[73]

Cost-effectiveness of donepezil hydrochloride

There are two grounds for prescribing a new drug. Firstly, it may obtain new health gains that are considered worth the additional cost (i.e. it is cost-effective in some broad sense), and/or secondly it lowers other health-service or broader social costs (this occasionally may happen to such an extent that a drug can be, in total, cost neutral or cost saving). There is currently no evidence to evaluate whether either of these objectives can be adequately met by donepezil hydrochloride. To date it is known that donepezil hydrochloride improves cognitive functioning over a short period, but the value of this to patients and their carers, and the consequences of long-term treatment are all unknown. Similarly, the effect of drug treatment on the broader cost of illness of Alzheimer's disease is currently unknown. A useful starting point, in the face of such uncertain benefits, is to ask under what circumstances donepezil hydrochloride would be cost neutral.

Donepezil hydrochloride (Aricept®) is reimbursed by the NHS at £2.44 for a 5-mg daily dose, and £3.42 for a 10-mg daily dose. Thus the annual purchase cost of treatment is £890 at the 5-mg dose or £1250 at the higher dose. Assuming one-quarter of patients require the higher dose, the average purchase cost is £980 per year. This may be set against current annual NHS care estimated to cost (on average) £3560 per patient: drug intervention (assuming it was permanent) would have to (on average) reduce by 28% these associated costs of illness of Alzheimer's disease to be cost neutral to the NHS. However this is too simplistic, and a number of uncertainties remain even for this simple presentation.

- Only mild to moderate dementia is indicated for treatment. Patients with this level of disease may make lower than average use of resources, making savings more difficult to achieve.
- Consultations or hospital admissions may be delayed by treatment rather than prevented.
- Constraining drug treatment to the exclusive and appropriate treatment of Alzheimer's disease may prove difficult in practice.
- The impact of treatment in the earlier stages of the disease on the later and more expensive stages is unknown. Longer-term follow-up may demonstrate worthwhile health gains and/or health-service savings, although there is currently no evidence for these effects.
- Consideration of (reductions in) broader social costs may make treatment more attractive, although there is currently no evidence to support this.

The average GP pharmaceutical budget is about £120,000 per year. To treat one patient with this one condition involves committing nearly 1% of the budget to 0.05% of patients. This investment has not been shown to be justified by an important health gain and cannot be currently advocated in a climate of scarce resources. It remains the responsibility of the company developing donepezil hydrochloride (and future competitors) to demonstrate the value of their product in terms that will demonstrate its worth.

Comment

In this guideline the group were looking at the introduction of a novel therapy for which there was only short-term evidence of effectiveness (but around which there was a high-profile marketing campaign). The group concluded that whether or not donepezil was a worthwhile treatment for Alzheimer's disease was not established by the currently conducted trials. However, it was clear that the drug had an effect upon cognitive function, and further longer-term randomised trials were required to evaluate the benefits and costs of donepezil in clinical practice in the UK or similar health system. Thus, the recommendations were driven by concerns about the effectiveness and safety of the drug rather than by cost issues. As a consequence the group made the following recommendation.

- In the light of limited current knowledge, GPs should not initiate treatment with donepezil (Aricept) nor continue hospital-initiated treatment (A).

The early management of schizophrenia: pharmacological treatments

The aim of this guideline[106] was to provide recommendations to guide healthcare professionals in the appropriate use of anti-psychotic drugs for the early management of patients with schizophrenia. There was considerable uncertainty about the comparative effectiveness of the new 'atypical' drugs (risperidone, sertindole, quetiapine, olanzapine, amisulpride and clozapine) in the treatment of schizophrenia and the guideline therefore concentrated on this issue. In trials, the atypical drugs are most commonly compared with two conventional neuroleptic drugs, haloperidol and chlorpromazine.

Trials of atypical anti-psychotics generally show considerable variability in efficacy and tolerability, when compared with conventional neuroleptic drugs, making simple combined estimates from trials of limited value. Analysis by drug suggests small benefits in reduced psychiatric symptoms favouring some anti-psychotics. Most trials are short term (6–8 weeks) and thus provide limited evidence on how best to treat patients in the longer term. There is no evidence of specific effects for atypical drugs upon negative and depressive symptoms. Effects, when they occur, seem equally to involve all classes of symptoms. Furthermore, there is inadequate information in direct randomised comparisons of atypical drugs to provide reliable evidence on their relative effectiveness. There is limited evidence of improved tolerability with olanzapine compared with risperidone.

Observed differences in the results of trials may be explained by variation in the dose of the comparator conventional neuroleptic used. This conclusion appears to be valid for controlled trials of both haloperidol and chlorpromazine. Once variability in dose is taken into account, the apparent benefits for the atypical anti-psychotics on overall symptom scores are no longer present, indicating that in trials conventional drugs are frequently used in doses that are inappropriately high. Drop-out rates are lower in groups treated with atypical anti-psychotics compared with those treated with conventional neuroleptics at high doses. However, in trials that use appropriate doses, the tolerability to conventional neuroleptics appears to be similar to that to the newer drugs, although the atypical anti-psychotics do appear to be associated with a reduced risk of extrapyramidal side-effects even in trials in which lower doses of haloperidol are used. There are no direct data on the relative incidence of tardive dyskinesia in trials of haloperidol at lower doses.

The costs and consequences of drug selection

Two studies formally conducted 'within trial' economic analyses (no extrapolation of cost or outcome was made beyond the follow-up of the trial). Both quantify the use of resources (*Tables 11* and *12*).

In both studies, patients were allowed to switch therapy at any time if required, and hence the reported use of resources is pragmatic. In the study by Rosenheck and colleagues[107] clozapine led to reductions in hospitalisation, whereas in the Risperidone Outcome Study of Effectiveness (ROSE)[108] risperidone did not. Greater use of outpatients services by patients receiving clozapine

TABLE 11 *Use of resources, at 1 year, for patients treated with clozapine or haloperidol (Rosenheck,* et al., *1997[107])*

Variable	Clozapine	Haloperidol	*p* value
Inpatient or residential care (days)	158.7	179.8	0.07
Psychiatric	143.8	168.1	0.03
Medical, surgical or other	14.9	11.7	0.5
Inpatient psychiatric re-admissions	1.7	1.5	0.22
Outpatient services (units)	133.6	97.9	0.03

TABLE 12 *Use of resources, at 1 year, for patients treated with risperidone or conventional neuroleptics in the ROSE study (Meredith,* et al., *1998[108])*

Variable	Risperidone	Conventional neuroleptics
Index hospital days	10.5	8.8
Acute hospital (non-index)	19.6	18.7
Partial hospitalisation	12.1	14.4
Emergency room	0.8	1.1
Crisis team/crisis bed	1.4	1.8
Total acute care days	43.4	43.5
Routine psychiatric care (visits)	31.3	28.4

is likely to be underestimated since patients in the haloperidol group attended for the blood tests required by patients receiving clozapine to monitor the incidence of agranulocytosis. Rosenheck and colleagues adjusted for this effect in their subsequent costings. Both studies applied US healthcare costs to estimate the net impact on cost of drug selection (*Tables 13* and *14*). Despite reasonable numbers randomised in both studies, the overall effect upon net costs remains uncertain. Rosenheck and colleagues suggest net healthcare cost savings for clozapine due to reduced hospitalisation, but the 95% CI is wide: –$9250 to $3780 (estimated from the *p* value using a *t* distribution). The ROSE study reports, on average, an increase in costs mainly due to the higher cost of risperidone itself. Again, however, there is considerable uncertainty (95% CI for net cost, –$1828 to $5755).

The main difficulty with both of these analyses is interpreting what they might mean in a UK healthcare context. Unit costs of hospitalisation are generally considerably higher in the US setting, and so any savings from reduced hospitalisation would look less impressive in the NHS context. Additionally, it is unclear to what extent the use of resources themselves might differ: for example, UK psychiatric patients may have longer or shorter average hospital stays.

Cost per case of tardive dyskinesia avoided

The reduction of tardive dyskinesia when using olanzapine instead of haloperidol can be estimated

TABLE 13 *Net cost of care, at 1 year, for patients treated with clozapine or haloperidol (Rosenheck,* et al., *1997[107])*

	Cost (US$) for patients treated with:		
Variable	Clozapine (C_C)	Haloperidol (C_H)	*p* value ($C_C - C_H$)
Inpatient or residential	49,311	56,752	0.03
Psychiatric	45,247	53,931	0.01
Medical, surgical or other	4,064	2,821	0.2
Outpatient	8,473	3,474	< 0.001
Patient care	5,274	3,107	< 0.001
Anti-psychotic medication	3,199	367	< 0.001
Total healthcare costs	57,785	60,226	0.39
Total non-healthcare costs*	366	659	0.008
Total cost to society	58,151	60,885	0.41

* *Productivity, criminal justice, family burden (lost income), transfer payments*

TABLE 14 *Net cost of care, at 1 year, for patients treated with risperidone or conventional neuroleptics in the ROSE study (Meredith,* et al., *1998*[108]*)*

Variable	Cost (US$) for patients treated with:		
	Risperidone (C_R)	Conventional neuroleptics (C_C)	95% CI* ($C_R - C_C$)
Index hospital days	6,035	5,057	–702 to 2656
Acute hospital (non-index)	11,255	10,757	–2491 to 3485
Partial hospitalisation	2,217	2,776	–1586 to 467
Emergency room	112	146	–92 to 25
Crisis team/crisis bed	433	544	–379 to 157
Total acute care days	20,055	19,284	–2962 to 4502
Routine psychiatric care (visits)	959	875	–72 to 239
Estimated drug costs	2,695	1,586	811 to 1407
Total estimated costs	23,709	21,746	–1828 to 5755

* *Wilcoxon tests of differences between costs were used by the authors to estimate CIs*

from one published study.[109] Its limitations include the use of relatively high doses of haloperidol and long-term patients with chronic schizophrenia, in whom side-effects are most prevalent. Notwithstanding the concerns about the validity and generalisability of this evidence, it is possible to estimate the cost per case of irreversible tardive dyskinesia avoided. Such an approach is necessarily tentative, but may provide some indication of the resource consequences of attempting to reduce this important side-effect.

Based upon the average annual cost of typical and atypical anti-psychotics in the current mix of prescribing, and the reduction of risk observed in the study of Tollefson and colleagues,[109] the cost per case of tardive dyskinesia avoided is estimated to be £30,600 (95% CI, £14,600 to £88,800) (see *Table 15*). However, the extent to which this represents a general cost for avoiding tardive dyskinesia is unclear given the uncertainties concerning: the definition of tardive dyskinesia; the impact on different patient groups; the effect of the dose of haloperidol; the effect with alternative conventional drug; and the effect with alternative atypical anti-psychotic drug.

Budgetary implications of changing to atypical anti-psychotics

In 1997 only about 6.2% of the volume of prescribing of anti-psychotic drugs included atypical anti-psychotics. In this section we explore the budgetary implications for a typical health authority of different scenarios for switching to

TABLE 15 *Cost per case of tardive dyskinesia avoided in the study reported by Tollefson,* et al. *(1997)*[109]

	TD rate (%, 220 days)*	TD rate (%, 1 year)	Cost (£,1 year)†	Cost (£) per case of TD averted
Olanzapine	0.99	1.64	1983	–
Haloperidol	4.57	7.58	163	–
Olanzapine – haloperidol (95% CI)	3.58 (1.24 to 7.51)	5.94 (2.05 to 12.46)	1820 (–)	30,600 (14,600 to 88,800)

* *Treatment durations reported were similar: olanzapine, median 237 days; haloperidol, median 203 days. For the purpose of estimation the average of the medians was applied to both groups*

† *Mean reported end-point doses (olanzapine 14.41 mg; haloperidol, 14.67 mg) are used to estimate cost. Costs are derived from average prescribing costs and quantity data for English primary care in 1997, and thus reflect the mix of forms in which the drugs are currently used. It is assumed that there is no net change in costs other than the cost of the drug itself. (The validity of this assumption is unclear given the various uncertainties in the estimate – see text)*

TD, tardive dyskinesia

TABLE 16 *Scenarios exploring the implication for English NHS primary care prescribing costs of increased use of atypical anti-psychotics*

Scenario	Usage split (%)	Patient-years (1997)	Cost per year (£,1997)	Cost (£, 1997)
Current pattern (1997)				
Atypical	6.2	12,303	1674	20,600,260
Typical	93.8	186,341	97	18,045,206
Total	*100*	*198,643*	–	*38,645,466*
Low guess				
Atypical	21.0	41,715	1674	69,849,752
Typical	79.0	156,928	97	15,196,914
Total	*100*	*198,643*	–	*85,046,666*
High guess				
Atypical	62.0	123,159	1674	206,223,077
Typical	38.0	75,484	97	7,309,908
Total	*100*	*198,643*	–	*213,532,985*

Scenarios assume that the relative proportion and mix of drugs used within the typical and atypical groups remains constant. Treatments are assumed to be prescribed at WHO DDD

atypical anti-psychotics. Overall drug prescribing rates are adjusted using defined daily doses to provide volumes of patient treatment. In 1997 the total cost of anti-psychotics prescribed in primary care was £38.6 million (*Table 16*). Dividing this figure by 100 provides an approximate average estimate for one health authority of £0.39 million.

Given the relatively high cost of newer drugs, overall cost is sensitive to quite small changes in prescribing behaviour. In the ROSE trial,[108] patients were randomised to receive risperidone or conventional therapy, but substantial switching of drugs occurred in both groups. Thus the trial provides estimates of the extent to which patients on conventional therapy may be switched to atypical drugs by psychiatrists accustomed to their use. It also estimates the extent to which patients treated initially with an atypical anti-psychotic may be switched to conventional therapy because of poor response or side-effects.

Among patients randomised to conventional care, 79% of treatment-days were provided by a conventional neuroleptic and 21% by risperidone. Among patients randomised to risperidone, the proportions were 62% for risperidone and 38% for conventional neuroleptic. These proportions are used to explore the possible prescribing implications of first-line prescribing strategies with typical and atypical drugs. Taking the ROSE[108] conventional care group data as a proxy for a conservative increase in the use of atypical drugs suggests anti-psychotic prescribing costs in primary care may rise to £85 million: an increase of £46 million or about £0.46 million in an average size health authority. A more sweeping move towards the use of atypicals, proxied by the ROSE risperidone group drug use, suggests primary care prescribing costs may rise to £214 million, an increase of £175 million or about £1.75 million in an average size health authority.

These costs are speculative, and are calculated on the basis of available evidence on the likely rates of usage of atypical drugs. Review work undertaken for this guideline has provided no conclusive evidence of desirable rates of prescribing of the newer drugs. Newer drugs vary substantially in price (*Figure 1*) and current prescribing indicates an increasing use of olanzapine, the most costly atypical anti-psychotic. There are inadequate data to explore the relative cost-effectiveness of the various atypical anti-psychotic drugs.

Other economic analyses

In addition to the within-trial economic analyses reported in this guideline a number of other published economic analyses were retrieved. These involved modelling of data from selections of various trials reported in this guideline and data from other sources. One such analysis[110] merits comment because it was produced by the Development and Evaluation Service, an NHS regionally funded body. The report explores the role of olanzapine as first- and second-choice treatment for schizophrenia. The conclusions are based on a decision analysis model extrapolating to 1 year from the four published 6-week olanzapine trials, and include a range

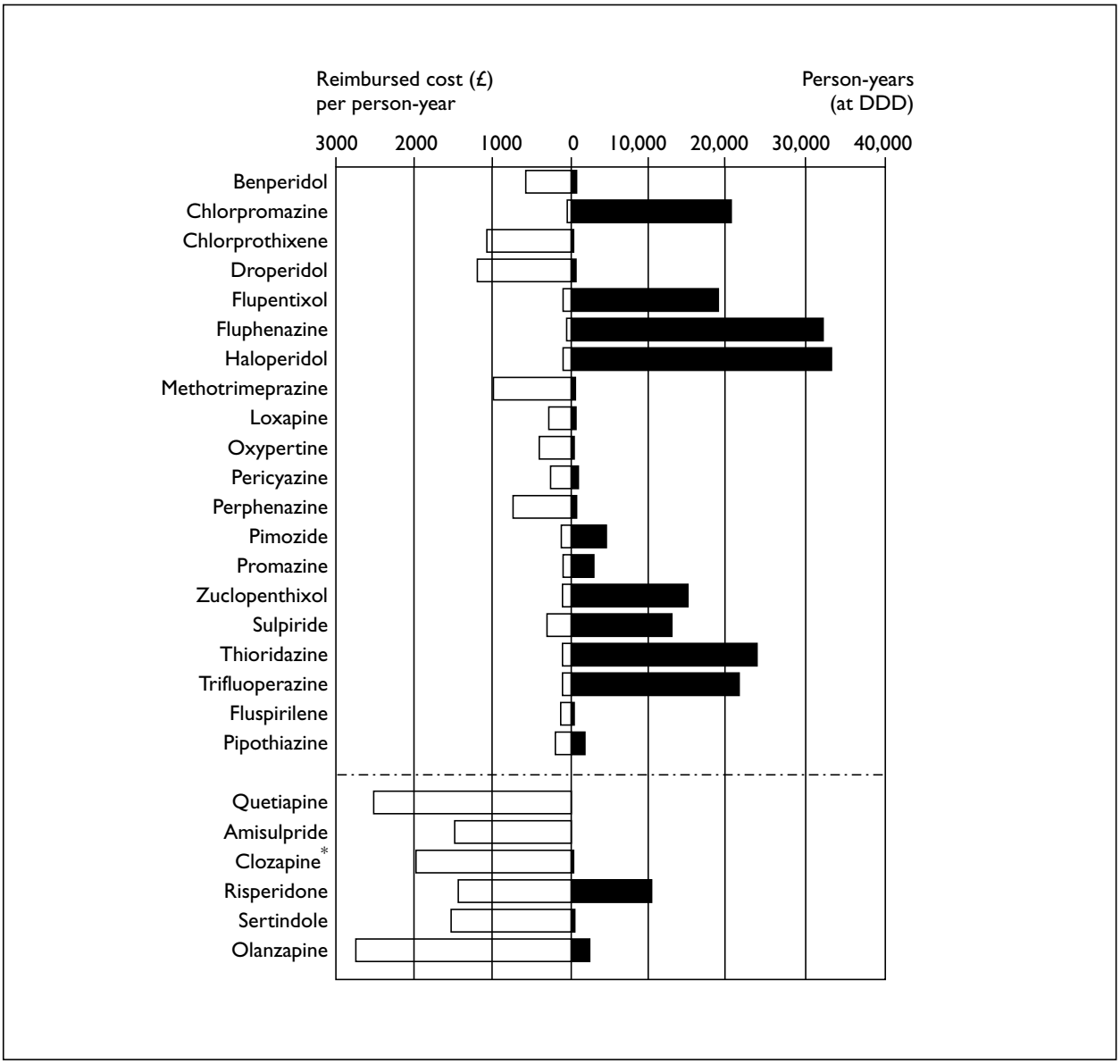


FIGURE 1 *Volume of use (■) and reimbursed cost per person-year of treatment at DDD (□) (England 1997)*

* *Includes cost of monitoring*

of quality-of-life and resource-use assumptions. From the guideline overview, it was not possible to validate most of the assumptions made in the model or the 'good evidence of excellent value for money' claimed for olanzapine. In particular, the claim of a robust cost advantage for olanzapine on the available evidence could not be substantiated. Evidence of the cost-effectiveness of atypical anti-psychotics from long-term randomised trials, reported in this guideline, does not support these assertions.

Comment

As was the case with donepezil, the evidence was inadequate to allow a properly informative summary of the relative costs and benefits of the new atypical anti-psychotic drugs. However, on the basis of the available evidence, it seemed likely that relief of symptoms for patients receiving atypical antipsychotics would be similar to (or for the most commonly used drugs, risperidone and olanzapine, possible slightly better than) that for patients receiving conventional antipsychotic drugs. However, it was unclear how much the apparent improvement with risperidone or olanzapine is, in fact, due to inappropriately high dosing of patients receiving conventional anti-psychotic drugs in trials. This theme recurred when trying to understand data on quality of life, resources and costs. Here, modelling the effect of dose suggested that differences in drop-out from treatment and efficacy

largely disappeared, although there was evidence that extrapyramidal symptoms may still be worse even with appropriate doses of conventional anti-psychotics. Overall, acceptability of the two drug groups may be more similar than available data suggest since atypical anti-psychotic drugs have their own side-effects, which are inconsistently reported in studies.

The group's overall conclusions were that the two studies presenting net costs of care show imprecise findings. The net cost of prescribing atypical anti-psychotic drugs in the UK setting remained uncertain and it could not be presumed that any savings from reduced hospitalisation or use of other services would offset the higher acquisition costs of the drugs. However, analysis based on the extension phases of three randomised trials comparing olanzapine and haloperidol suggests a significant reduction in tardive dyskinesia in chronic patients at high risk. Consequently the group made the following recommendation.

- Available (limited) data suggest that atypical drugs provide an acceptably cost-effective alternative for patients receiving conventional neuroleptics at appropriate doses who experience unacceptable extrapyramidal side-effects (in particular tardive dyskinesia) (B).

Chapter 6

Case study of a guideline that uses quantitative evidence summary methods and looks at a broad clinical area

One guideline was developed using quantitative summary methods and addressing multiple treatments in a broad clinical area for which a variety of treatments were available.

The primary care management of patients who have experienced a myocardial infarction

The aim of this guideline was to provide recommendations on the appropriate primary care management of patients who have previously experienced a myocardial infarction. The guideline focuses on the secondary prevention of disease, rather than the management of symptoms of angina, which is addressed elsewhere.[55] It sought to answer the (complex) question 'What are the benefits in mortality and major morbidity in identifiable major subgroups of patients who have experienced a myocardial infarction from treatment with statins, ACE inhibitors, antiplatelet agents, beta-blockers, calcium-channel blockers, cardiac rehabilitation, Mediterranean diet or polyunsaturated fatty acids and potassium-channel activators?' Specifically the guideline sought to examine and present the evidence concerning the appropriate sequencing of drugs in patients with a prior myocardial infarction and to identify whether this differs according to prognostic risk factors. Incremental health benefits and healthcare costs, for each treatment initiated, were presented as the available data permitted.

The effectiveness data showed that all the interventions under consideration (with the exception of potassium-channel activators) were effective in terms of mortality.

Valuing alternative treatments

For a GP and a patient trying to make a decision about appropriate long-term therapy following myocardial infarction the problem posed is: how should treatments be sequenced or combined? The recommendations of this guideline are intended to inform such decisions. However, the data from trials present two fundamental problems.

Firstly, patients included in the trials of different drugs are at different levels of underlying risk of further fatal ischaemic heart disease (approximated as the mortality rates found in trial control groups). Thus for two different drugs achieving the same relative reduction in mortality during the same period of time, one drug may appear to generate greater survival benefits although both may be equally effective. It is inappropriate when trying to quantify benefits of different drugs to a new patient to compare the trial-based absolute benefits of improved survival since these are obtained from different patients at different underlying risk.

Secondly, the findings of trials already have a drug sequence implicit in them. In major trials of beta-blockers the most common co-therapies are diuretics, nitrates and digitalis. A similar pattern of co-therapy (although inconsistently reported) is suggested for the trials of antiplatelet drugs, which were conducted during the same period. Hence both beta-blockers and antiplatelet drugs in patients following myocardial infarction can be thought of as preventative therapies as well as part of medical management. In major trials of the statins, the majority of patients received aspirin and substantial proportions of patients received beta-blockers, ACE inhibitors, calcium-channel blockers or some combination. A similar pattern emerges for the use of ACE inhibitors in patients at raised cardiovascular risk in the Heart Outcomes Prevention Evaluation (HOPE) trial.[111] Consequently findings for ACE inhibitors and statins in patients after myocardial infarction are compared with medical management including the use of beta-blockers and aspirin. It is unclear how large the benefits might be of giving a patient a statin or ACE inhibitor instead of a beta-blocker since such a head-to-head trial does not exist. It is more appropriate to consider adding treatments in the order that is implicit in the trials.

For patients with previous myocardial infarction and heart failure, virtually all patients in trials of beta-blockers received an ACE inhibitor. Only a minority of patients in trials of ACE inhibitors received a beta-blocker. In the one trial to date of spironolactone, nearly all patients received an ACE inhibitor but few received a beta-blocker. Thus the trials present the value of using an ACE inhibitor in patients with previous myocardial infarction and heart failure, and of adding either a beta-blocker or spironolactone to this.

Comparison between drugs is complicated further by differences in the type of disease and the point in the disease process at which patients are enrolled. To explore the effect of beta-blockers it has been necessary to analyse separately post-acute phase short-term trials and long-term follow-up trials. Trials of ACE inhibitors can be grouped into four categories.

1. Trials of (unselected) patients with previous myocardial infarction (who may or may not have heart failure). The major trials are short-term, starting immediately after an acute episode.
2. One long-term trial of patients at raised cardiovascular risk (HOPE). Most patients had a history of cardiovascular disease, about half of patients had a previous myocardial infarction and none had heart failure.
3. Trials in selected patients with prior myocardial infarction and heart failure.
4. Trials in patients with heart failure some of whom may have had a previous myocardial infarction.

Trials in all of these categories have been presented since they help to explore the effect of treatment. However, in the sections that follow, category 2 is used to best represent patients with previous myocardial infarction, and category 3 is used to best represent patients who additionally have heart failure. Trials of spironolactone have only been carried out in patients with severe heart failure, without specific reference to myocardial infarction, and so provide less direct evidence than that available for ACE inhibitors. Trials of beta-blockers have been carried out in patients with heart failure who may or may not have experienced myocardial infarction. Evidence from trials of beta-blockers in uncomplicated myocardial infarction enhances the plausibility that patients with both myocardial infarction and heart failure benefit as do those with heart failure. Calcium-channel blockers and potassium-channel blockers are not considered further since their ability to improve survival is not established.

Given that the trials do not provide the simple head-to-head results in comparable patients that might be hoped for, one response is to provide a profile of the important attributes of treatments (in this instance effectiveness, tolerability, and cost). The interpretation reflects the characteristics of the trials and seeks clear indications about how to use available treatments.

The profile approach

Mortality rates are measured consistently in trials and a rigorous analysis of the propensity of the different treatments to improve survival has been possible. One measure of co-morbidity that is also measured reasonably consistently in trials is non-fatal myocardial infarction and it is notable that this outcome broadly correlates with mortality data, reflecting an influence of treatments upon the underlying process of ischaemic heart disease (*Table 17*). Other aspects of treatments that are important to patients include tolerability, side-effect profiles and influence upon quality of life. These other aspects of treatment are less consistently reported in trials.

It is apparent that the various treatments offer similar improvements in survival in relative terms (*Table 17*). The absolute benefits (expressed as incident rate differences) are very different reflecting different baseline risk. It should not generally be presumed that the relative reduction in mortality achieved by a drug in patients at one level of risk would be achieved by patients at another level of risk. However, there is some evidence to support this assumption in drug treatments following myocardial infarction. Stratification of results by prognostic risk markers in the largest trials[111–114] provides evidence of similar relative reductions in mortality across risk strata. A cautious interpretation of the trials is that the drugs demonstrate similar effectiveness, albeit in different patient groups, and are similarly well tolerated. It is likely but not certain that all of beta-blockers, antiplatelet drugs, ACE inhibitors and statins will work (on average) similarly well in individual patients with previous myocardial infarction with the following two major caveats.

- Exclusions from trials of statins were extensive with the consequence that nothing is known about how well these drugs work in patients at higher risk.
- To reflect the sequencing of drugs in trials, the values for statins and ACE inhibitors should be considered as benefits in addition to those from appropriate first-line use of antiplatelet drugs and beta-blockers.

TABLE 17 *Comparison of summary findings from trials*

	Reduction in all-cause mortality: 1 – odds ratio* (%)	Reduction in all-cause mortality: incident rate difference†	Reduction in non-fatal MI: incident rate difference†	Reduction in non-fatal stroke: incident rate difference†	Withdrawal from treatment: % above placebo
Patients with MI					
Beta-blockers	24 (17 to 30)	13 (7 to 18)	8 (2 to 14)	NR	1.2 (0.6 to 1.8)
ACE inhibitors	17 (5 to 27)	4 (1 to 6)	5 (2 to 7)	3 (1 to 5)	1.7 (–0.2 to 3.5)
Statins	24 (5 to 40)	4 (2 to 6)	6 (2 to 10)	2 (1 to 3)	NR
Antiplatelet drugs	16 (2 to 27)	7 (1 to 13)	8 (5 to 11)	2 (1 to 4)	NS
Cardiac rehabilitation	26 (11 to 38)	9 (3 to 16)	NR	NR	NR
Mediterranean diet	21 (6 to 35)	6 (1 to 10)	NS	NS	NR
Patients with MI and diabetes					
Insulin	36 (11 to 55)	31(2 to 60)	NR	NR	NR
Patients with MI and HF					
Beta-blockers	35 (25 to 45)	35 (23 to 46)	NR	NR	5 (–19 to 30)
ACE inhibitors	26 (14 to 34)	18 (8 to 28)	NS	NR	NR
Spironolactone	38 (24 to 49)	57 (26 to 87)	NS	NS	5 (2 to 9)

* *The odds ratios presented are random effects estimates; 95% CIs are shown in brackets*

† *Incident rates are calculated as the reduction in the number of events for 1000 patients treated for 1 year when comparing treatment with placebo (see text for explanation)*

NR, not reported consistently; NS, data not reported consistently or in a form permitting quantitative summary, but indicating no significant difference between treatment and placebo

For patients with heart failure in addition to previous myocardial infarction it is possible to say that ACE inhibitors, beta-blockers and spironolactone offer worthwhile benefits with the following caveats.

- Spironolactone has only been tested in trials involving patients with severe heart failure. Benefits in patients with less severe heart failure, and in patients with myocardial infarction in particular, are unknown.
- To reflect the sequencing of drugs in trials, the values for beta-blockers and spironolactone should be considered as benefits in addition to those from appropriate first-line use of ACE inhibitors.

No useful quality-of-life data have been identified from trials relating directly to patients who have previously experienced a myocardial infarction.

Since resources are limited, the additional benefits of each new treatment should be set against the costs. Some major trials for some drugs report reductions in hospitalisations during the trial follow-up period: subsequent economic analyses offset the costs of the drugs with savings due to reductions in admissions. Unfortunately these savings are not reported consistently for different drugs. One approach might be to offset drug cost only when trial data are available to demonstrate this. However, savings achieved are affected by the underlying risk of patients enrolled in a trial in the same manner as mortality data. The various drugs are predicted to achieve similar relative reductions in mortality and morbidity and it is plausible that reductions in hospitalisations will be similar for all of the drugs for new patients being treated. A potential criticism of analysing reductions in hospitalisations measured in trials is that morbidity may have been delayed during the trial period but that patients still have a chronic disease that will progress in time. Additionally, although a drug may achieve additional survival, healthcare costs tend to increase with age. Consequently it is uncertain whether resource savings achieved in trials of cardiovascular disease are permanent or only temporary. Analysis restricted to using comparable data for all treatments means using just the acquisition costs of the drugs. It is recognised that the net cost of treatments may be overestimated by leaving out reductions in hospitalisation, but that this is likely to be a consistent bias that will not alter their relative costs.

All drugs included in major trials with a licensed indication for use in England have been costed. Drug treatments for patients with previous myocardial infarction and previous myocardial infarction with heart failure were costed using the dosing schedules found in the trials (shown as appendices in the full guideline report). Assuming a common class effect for the effectiveness and tolerability of drugs, the cheapest drugs are shown in *Table 18*. It is recognised that there is a potential impact on compliance when trading off cost against frequency of dosing. To reflect this, if the cheapest drug in the class requires frequent dosing, the cheapest drugs with once or twice daily dosing are also shown in *Table 18*.

Subcutaneous insulin infusion, as a prophylactic treatment for patients with diabetes and prior myocardial infarction, is not a licensed indication in England. Mediterranean diet-type interventions are problematic since patients may bear considerable costs in dietary modification and the trial interventions involved considerable investment by the trialists for which the required resources are unclear. Both of these interventions have been left uncosted although both interventions appear effective. National published data provide a cost for (non-elective) cardiac rehabilitation of £3035 (interquartile range, £1705 to £3503; HRG Code S23, 1998 data).

Considering the acquisition costs of drugs for patients with previous myocardial infarction it is clear that beta-blockers and aspirin are not only effective first-line treatment but also very good value for money. Similarly, all treatments for patients with both myocardial infarction and heart failure appear good value. In the light of the apparent survival gains, none of the available treatments shown present an unacceptable cost.

In summary, GPs and patients are presented with an array of effective drugs for patients after myocardial infarction, with or without complicating heart failure. The value of these should be discussed with patients, and if appropriate, initiated and continued if tolerated.

Limitations of the profile approach

A criticism of limiting the discussion to the effects measured during trials is that they may not describe benefits helpfully to patients. For example, patients may be interested to know how much longer they are likely to live by taking a (primarily) preventative treatment. Survival gains are considerably truncated by considering only

TABLE 18 *Comparative cost of drug treatments for patients with previous MI*

Drug class	Drug*	Daily dose†	Cost/year‡
Patients with MI			
ACE inhibitors	Ramipril (Tritace®)	5 mg b.d.	249
Antiplatelet drugs	Aspirin (generic)	75 mg o.d.	2
Beta-blockers	Propranolol (generic)	80 mg t.d.s.	8
	Metoprolol (Betaloc®)	100 mg b.d.	45
Statins	Pravastatin (Lipostat®)	40 mg o.d.	387
	Simvastatin (Zocor®)	20–40 mg o.d.	387
Patients with MI and HF			
ACE inhibitors	Captopril (generic)	25–50 mg t.d.s.	36
	Ramipril (Tritace®)	2.5–5 mg b.d.	241
Beta-blockers	Propranolol (generic)	40 mg t.d.s.	4
	Bisoprolol (Emcor®/ Monocor®)	5–10 mg o.d.	118
Spironolactone	Spironolactone (generic)	25 mg o.d.	22

* *For a full listing of the doses and costs of drugs, by drug class, with trial evidence for treatment in patients with MI or MI and HF see appendices in guideline report*

† *Doses shown are the target doses used in trials. Where a dose range has been reported the cost has been calculated by weighting the proportions of each dose used (see appendices in guideline report). Where the cheapest form of a drug represents the most convenient (once daily) dosing, this is listed exclusively. Where there is a trade-off between more convenient dosing and cost, several forms are listed*

‡ *Reimbursed cost to prescribe the cheapest priced form of the drug for 1 year. Source of cost (eMIMS, November 1999)*

the period of trial follow-up. Furthermore in trying to summarise the costs and benefits of different treatments in a profile we may in fact be running an internal and implicit modelling process in which we weigh up alternatives and try to make 'sensible' recommendations. This process may become strained when the available trials have failed to deliver simple and interpretable comparisons, as is the case for this guideline. Advocates of the modelling approach argue that it is best to use explicit assumptions to put the evidence together to explore value for money and to try to explore thoroughly the assumptions made to see if findings are robust. A modelling approach is presented and its limitations explored.

The modelling approach

The purpose of modelling is to help make meaningful comparisons between available treatments. This involves extrapolating survival gains over the remaining life-expectancy of patients. This is achieved by taking the survival rates from the control groups of trials (untreated patients) and modelling their continued survival, adjusted for age, as a hypothetical cohort of patients until all are dead. In the analyses that follow, the modelled treatment group is similar to the control group with the exception that survival is improved for a treatment period of 5 years, reflecting the relative risk reduction seen in trials. That is, the decision to provide a treatment for the next 5 years is being explored. For example, the modelled survival curves from trials of ACE inhibitors including patients with previous myocardial infarction and with or without heart failure are shown in *Figure 2*. For reference the population average survival curve is shown.

The gain in survival is the area between the treatment and control survival curves. It is notable that patients with heart failure are dying very much more quickly than those without, reflecting more severe underlying disease. Consequently nearly half of the gain attributable to treatment accrues in the modelled 5-year treatment period for patients with heart failure. In patients without heart failure about 80% of the estimated gain from treatment occurs after the 5-year treatment period.

Survival gains are similar for patients of different ages and sex: survival gains for men and women aged 65 years at the start of treatment are shown in *Figure 3*.

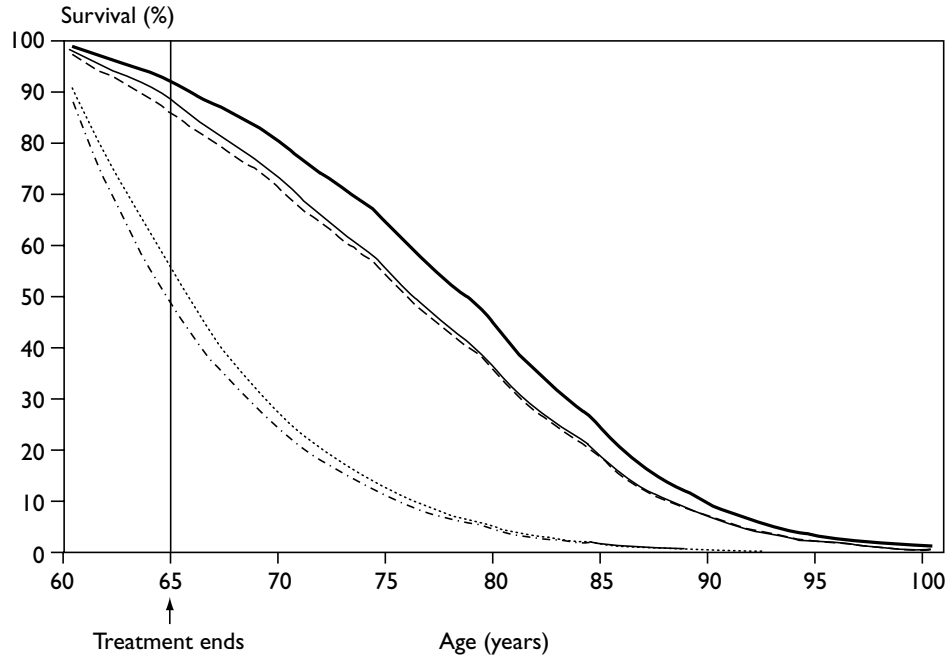


FIGURE 2 *Survival curves* extrapolated from trials of patients receiving an ACE inhibitor or placebo (——, population average; ——, post-MI with ACE inhibitor; – – –, post-MI without ACE inhibitor; ·····, post-MI + HF with ACE inhibitor; – · – ·, post-MI + HF without ACE inhibitor)*

* *Curves shown are for men aged 60 years at the start of treatment*

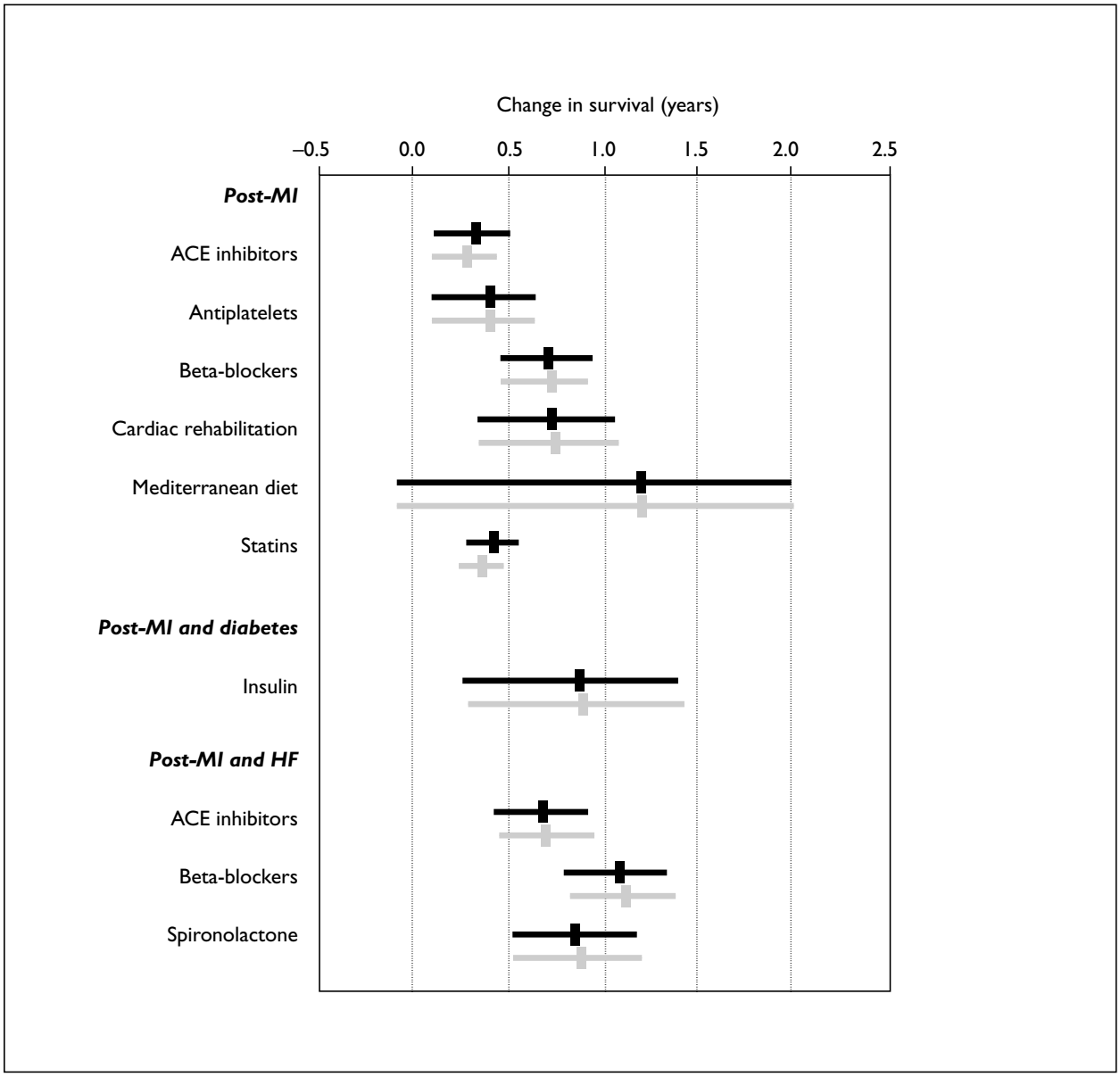


FIGURE 3 *Estimated survival gains (undiscounted) and 95% CIs associated with different treatments for patients with prior MI, alone or with diabetes or HF*[*]

[*] *Results shown are for men (▬) and women (▬) aged 65 years at the start of treatment. The vertical bar indicates the mean and the horizontal bars show the 95% CIs*

It is notable that predicted survival gains are more similar for the various drugs than are the corresponding incidence rate differences from the trials. The benefit predicted for dietary intervention is not statistically significant because of the cautious use of random effect estimates of relative risk in the modelling process. Treatments are predicted to extend average survival in a patient by between about one-quarter and 1 year of life.

For reasons identified in the previous section, a cautious approach to the cost of drug treatments would simply be to set gain in life-expectancy against the 5-year cost of prescribing each drug (or in the case of rehabilitation, a one-time referral). Costings are not available for Mediterranean diet or insulin interventions. For cardiac rehabilitation the mean cost for inpatient non-elective rehabilitation has been applied. The cost-effectiveness of treatments is shown in *Figure 4* for men starting treatment at age 65 years. (The findings were not sensitive to age or sex.)

Reflecting the wide range of costs of treatment, the range of cost-effectiveness of treatments

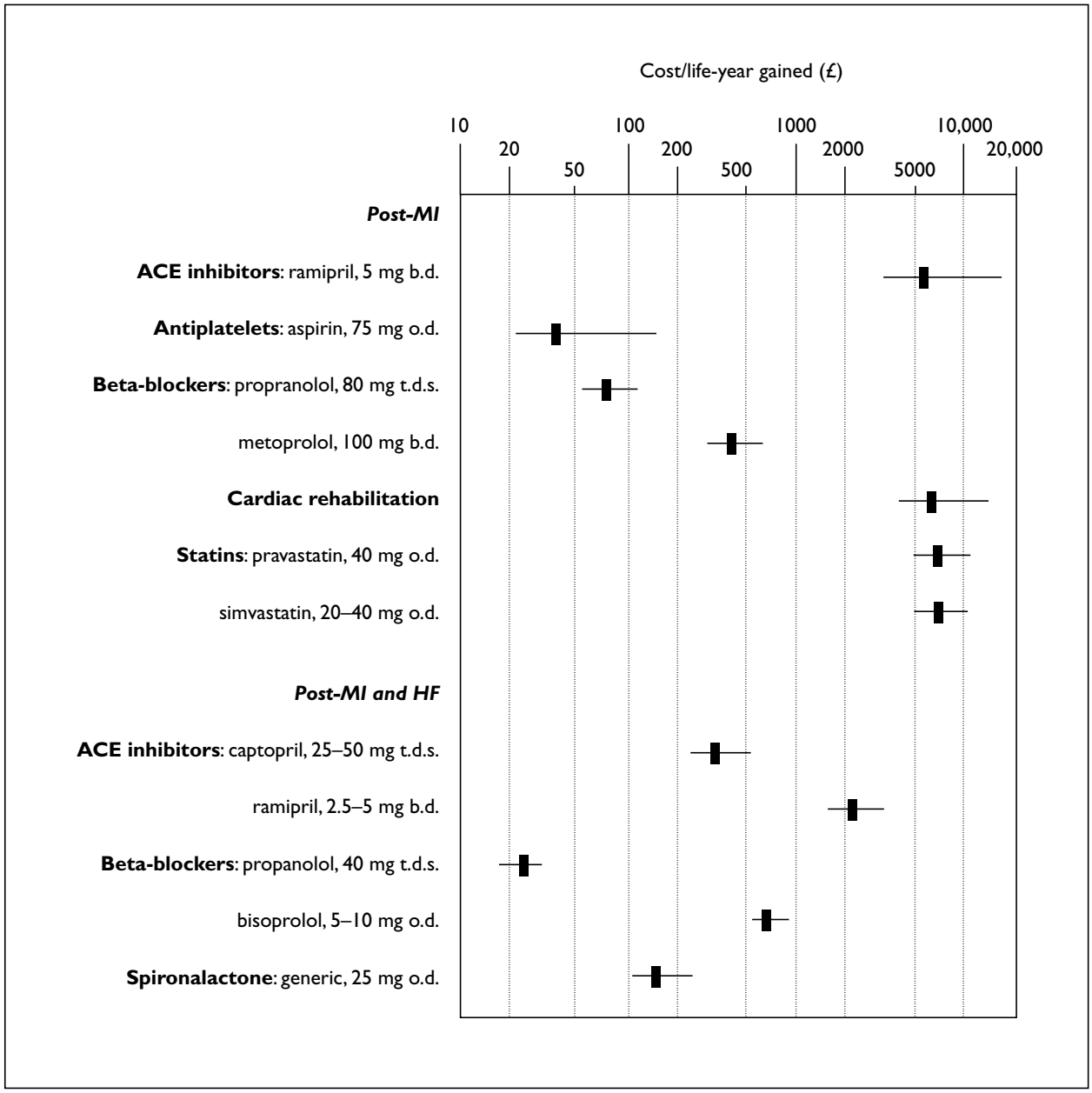


FIGURE 4 *Estimated cost-effectiveness and 95% CIs of different treatments for patients with prior MI, alone or with HF*[*]

[*] *Results shown are for men aged 65 years at the start of treatment. The vertical bar indicates the mean and the horizontal bars show the 95% CIs*

is quite broad. However, the method of estimating costs is conservative and it is apparent that all available treatments fall within accepted bounds of cost-effectiveness. The purpose of the modelling exercise is to provide comparable estimates of cost-effectiveness and *Figure 3* gives such a presentation **reflecting the patients included in the trials**. That is to say we do not know how these cost-effectiveness estimates might change when applied to a new patient with one level of underlying risk rather than the different ones found in the trials of the various treatments.

This can be explored by putting into the model a constant baseline risk for all the treatments and assuming the relative risk reduction for each treatment derived from the trials can be applied. There is some weak evidence to support this step, but nonetheless it is a strong modelling assumption. The consequence of running such a model is that the survival gains become more similar for each of the treatment alternatives (*Figure 5*).

This latter model explores the consequence of trying to adjust for different underlying risk in the various trials and suggests that the treatments

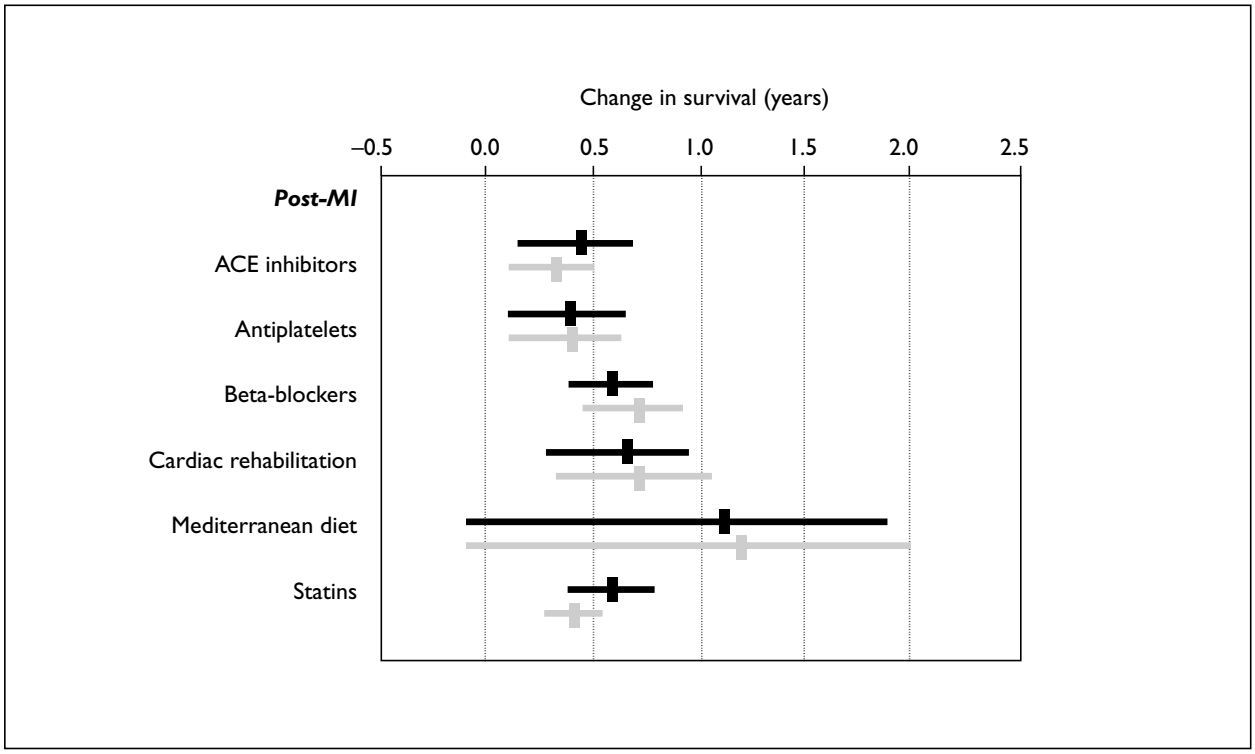


FIGURE 5 *Estimated survival gains and 95% CIs of different treatments for patients with prior MI[*], assuming a common underlying baseline risk for all treatments (adjusted baseline risk, ▬)[†] or an underlying baseline risk reported in trials (trial baseline risk, ▬]*

[*] *Results shown are for men aged 65 years at the start of treatment. The vertical bar indicates the mean and the horizontal bars show the 95% CIs*

[†] *The common underlying risk assumes an initial 5% annual all-cause mortality in all patients*

may in fact have similar effectiveness in terms of prolonging life in any particular risk group. Recalculated cost-effectiveness estimates using this pattern of survival are similar to those shown in *Figure 4* and qualitatively the same as estimates based more directly on the trials: this is because the cost-effectiveness estimates are most influenced by the costs of the drugs.

In summary, a modelling approach suggests all available treatments (in *Figure 3*) demonstrate worthwhile survival gains. Modelling additionally suggests that differences in the survival gains attributed to drugs from trials may be partly explained by differences in the underlying risk of enrolled patients. In patients with previous myocardial infarction, beta-blockers, antiplatelet drugs, ACE inhibitors and statins all appear to offer acceptable cost-effectiveness and should be initiated and continued by patients if tolerated, with the following two caveats.

- Exclusions from trials of statins were extensive with the consequence that little is known about how well these drugs work in patients at higher risk.
- To reflect the sequencing of drugs in trials, the values for statins and ACE inhibitors should be considered as benefits in addition to those of appropriate first-line use of antiplatelet drugs and beta-blockers.

For patients with heart failure in addition to previous myocardial infarction it is possible to say that ACE inhibitors, beta-blockers and spironolactone provide worthwhile survival gains and are cost-effective with the following caveats.

- Spironolactone has only been tested in trials involving patients with severe heart failure. Benefits in patients with less severe heart failure, and patients with myocardial infarction in particular, are unknown.
- To reflect the sequencing of drugs in trials, the value of beta-blockers and spironolactone should be considered as benefits in addition to those of appropriate first-line use of ACE inhibitors.

Limitations of the modelling approach

For trials reported in this guideline, follow-up is generally modest and therefore substantial extrapolation is required to estimate gains in

life-expectancy caused by treatment. When extrapolations are conducted these calculations are not value-free. The lower the underlying rate of mortality, the more substantial the extrapolations become and the more speculative the predicted gains from treatment. Patients who have survived a period of treatment on a particular drug are atypical, unlike healthy people or 'average' patients with the same medical condition. Although analysts strive for face validity and plausibility in their models, extrapolation of survival requires assumptions that cannot be validated.

Both a 'profile' approach and a 'modelling' approach were used to help explore the interpretation of trial findings. Although quite different in presentation and assumptions, both support the inference that when sequenced appropriately a range of treatments are both effective and cost-effective in patients with prior myocardial infarction.

Comment

The ambitious objectives set for the guideline led to considerable debate in the group about how best to describe and summarise the evidence. Attempting to sequence treatment options by comparing evidence derived from several meta-analyses of very different patient populations had not been attempted before. The modelling used in this guideline was more extensive than in any of the preceding ones but the process was again driven by the guideline development group's wish to explore the evidence in as robust a manner as possible. The guideline development group reached the following recommendations.

- Patients with prior myocardial infarction should receive long-term treatment firstly with a beta-blocker and aspirin, and then with a statin or an ACE inhibitor, reflecting the evidence from trials and estimates of cost-effectiveness (A).
- Patients with prior myocardial infarction and heart failure should be treated long-term with an ACE inhibitor and then a beta-blocker. In addition, patients who have moderate or severe heart failure (New York Heart Association grade 3 or 4) should be treated with spironolactone. All of these treatments are cost-effective (A).

Chapter 7

Case study of a guideline developed using decision analysis

One guideline development used a decision analysis model to provide the core of its evidence base and thus is distinct from other guidelines reported here.

Anticoagulation to prevent stroke in patients with atrial fibrillation

A Markov decision analysis was developed to explore the use of warfarin in patients with atrial fibrillation, using systematic literature review and appraisal, supplemented by additional research, to inform a guideline development group. As the conduct and findings of decision analysis differ in a number of respects from traditional evidence-based approaches, the development of this guideline is reported in more detail than the other case studies. However, this is still a summary of the full process and for a detailed understanding the reader is directed to the full report.[115,116]

Methods

A multidisciplinary guideline development group, which met on three occasions and included a cardiologist, a haematologist, a geriatrician and five GPs, helped to define the scope of the work and to develop explicit questions for literature review and modelling, and advised on the guidelines produced. Systematic literature searches were conducted to identify relevant papers on: effectiveness of anticoagulant and antiplatelet therapies; natural history and stroke risks of patients with atrial fibrillation; adverse effects of warfarin; utility of relevant health states; and costs of treatment. Only one identified study included relevant health state values, on a population thought unlikely to be representative of the UK. Therefore, a utility assessment exercise was conducted, using the standard gamble method. Health states were ranked and anchored to 'normal health' and 'immediate death' using standard methods. Fifty-seven elderly volunteers from a representative sample of community-based patients with atrial fibrillation identified in a previous study were interviewed to derive utility measures for relevant health states. NHS costs were calculated. Only one study allowed inpatient stroke costs to be broken down by severity. These were inflated to account for outpatient and primary care costs. Local cost data were derived from a hospital-based pharmacy-led anticoagulation service. The cost of a gastrointestinal bleed was based on mean length of stay and mean daily cost.

The decision model

The treatment decision was modelled as a Markov process using DATA 2.6 software. Over time, patients in the model move between several health states, reflecting the progression of disease. The model is run twice, with and without patients receiving warfarin, and thus the additional benefits of reducing morbidity over time can be explored. Data on effectiveness of warfarin, absolute risk of stroke, risk of recurrent stroke, outcome of stroke and risk of major (non-cerebral) bleed were derived from systematic review, with point estimates used in the model. As utility values were not normally distributed, median values were used.

All such models require a number of assumptions (*Box 15*): life-expectancy was taken from official statistics and a relative risk of 1.92 for all-cause mortality for patients with atrial fibrillation compared with the general population was applied. Adjustment was made to avoid double-counting fatal strokes.

BOX 15 Key assumptions in the decision analysis model

- The model covers remaining life-expectancy
- Patients are on warfarin for the first year only (this was considered to best match clinical practice, with review of treatment occurring at set intervals to take account of, for example, changing contraindications or clinical evidence)
- The relative risk reduction afforded by warfarin is constant across different absolute risks of stroke
- The relative risk reduction afforded by warfarin is constant across different severities of stroke
- The outcome of stroke is constant across different absolute risks of stroke
- Warfarin offers no protection against mortality from other causes
- Minor bleeds are not considered
- All events occur 6 months into the year

In keeping with clinical practice, the treatment decision was modelled for 12 months, after which time the decision is assumed to be re-assessed (i.e. the model is re-run using the updated information). The model was run for 1512 combinations of age, sex, blood pressure and risk factors and assessed results in terms of both QALYs and costs, both discounted at 5% per annum. Differences in QALYs and costs were calculated to determine which arm of the decision model (i.e. treatment with warfarin or not) maximised QALYs and minimised costs. The model outputs were sensitive to variation in a patient's utility for being on warfarin and to the estimate of warfarin effectiveness.

A number of assumptions were applied to the output from the decision analysis to produce a flow chart and look-up tables. The model could not determine what is an acceptable cost per QALY gained. Therefore, the guidelines were based purely on effectiveness in terms of QALY gain, given the low cost per QALY gained when compared with other commonly accepted interventions: for example the estimated cost per QALY gained from ACE inhibitor treatment of people with a diastolic blood pressure of 100 mmHg is between £11,000 and £39,000 depending upon age and sex.

A flow chart identified high-risk patients for whom anticoagulation could clearly be recommended. If the decision cannot be made from the flow chart, users are referred to tables (derived from the decision analysis) incorporating age, systolic blood pressure and risk factors, to support the patient decision on the basis of individual risk profiles.

There were four outcome possibilities from the model. In the first, treatment produced QALY gains and cost savings, leading to an unequivocal decision to 'definitely treat'. In the second, treatment led to both QALY losses and higher costs, and the unequivocal decision was 'definitely do not treat'. In the third, treatment yielded more QALYs than no treatment but with higher costs and the decision was to 'treat if the cost per QALY gained is acceptable'. In the fourth, there was a QALY loss from treatment but at lower cost (i.e. a cost per QALY gained by not treating), the implication being that treatment would save money but also lower the patient's quality of life; the decision was 'definitely do not treat'.

Using this classification, a set of twelve age/sex tables were constructed. In the great majority of cases modelled treatment led to lower costs – in only 12 of the total of 1512 cells (0.8%) was there a cost per QALY gained, ranging from £250 to £6000 using the basic assumptions in the model.

Guideline derivation

The guideline development group raised issues of heart-rate control, cardioversion, and contra-indications for warfarin. These were not included in the model, but were addressed by introductory statements, produced by group consensus. When using 'the basic model', the advice for all patients with three or more risk factors, and for men with left ventricular hypertrophy and any one other risk factor, was to treat with warfarin. This was incorporated in the flow chart. Otherwise, clinicians were referred to the associated tables.

The increased risk of cerebral bleed associated with warfarin was taken into account in the model, with the effectiveness of warfarin estimated for all strokes, both haemorrhagic and ischaemic. However, for patients with a baseline risk of stroke less than 50% greater than the risk of cerebral bleed on warfarin, the reduction in the risk of stroke afforded by warfarin could be outweighed by this increased risk (assuming warfarin affords an approximate two-thirds reduction). Using an analogous method to that used in estimating the risk of a non-cerebral bleed, the risk of cerebral bleed on warfarin was estimated: it ranged from 0.15% in patients aged 60 years to 1.6% in patients aged 85 years and over. This suggests that it may be prudent not to treat patients below risks ranging from 0.23% (0.15 × 1.5) at aged 60 to 2.4% (1.6 × 1.5) aged 85 and over. In only ten of the 1512 cells, all for women aged 80 years and over, did the results of the model based on median values for patient utilities recommend treatment for patients below these thresholds. These cells are hence classified as 'do not treat' in the tables.

Comment

The use of decision analysis with its explicit population of a model enabled the elements of the decision-making process, and their implications, to be made explicit. It also enabled incorporation of a wider range of available 'evidence', particularly patient utilities. While the approach allows for explicit quantification of the uncertainty that underlies an apparently straightforward binary clinical decision this was only explored for two dimensions of the decision – the uncertainty around the effectiveness of warfarin and patient utilities. In both circumstances the result of modelling such uncertainty

was considerable. Interestingly, it was not possible to model the impact of varying the two dimensions together.

The interaction between the guideline development group and the evidence (in this case the construction of the model) was very different from the interactions within the other case studies described. Once the clinical problem had been scoped there was little remaining role for the group and they were not called upon to discuss the evidence or the implications of the model. The model produces the decisions and users of the model are required to trust the product. Such packaging of data does not allow explicit consideration of the uncertainty around the various dimensions of the decision by those involved in making it. It is currently unclear how the model decisions relate to what a patient would decide. Nor is it clear that it is necessary (or feasible) to collect explicit patient utilities within routine care settings.

While a number of these uncertainties are clearly amenable to empirical research (reflecting the relative absence of attempts to use decision analysis in guideline development), the dissociation between the guideline development group and the synthesis of the evidence is a unique feature of this case study.

Chapter 8
Discussion

The rationale for the development of clinical practice guidelines is to present a rigorous exploration of the evidence and delivery issues surrounding treatment options in healthcare, conducted by appropriately constructed groups of health professionals, consumers and specialists.

We have presented our experiences based on a case series of 11 guidelines developed over a period of 5 years. The initial focus of this project was to explore the methods of incorporating cost issues within clinical guidelines. However, this exploration has been paralleled by a more general development of the methods of treating evidence within the process of developing guideline development. Therefore, the process of reviewing evidence in guideline development groups incorporates an increasing sophistication not only in considerations of cost but also in review techniques and group process. At the outset of the project it was unclear how narrowly or broadly the concept of 'cost' could be considered. It is now clear that, alongside the effectiveness data and data describing quality of life, cost issues can successfully be represented as part of a broad profile of treatment attributes.

The use of an epidemiological and health-service resource summary early on in the process of developing each guideline has proved a useful device to begin the process of thought in each guideline development group about the importance and value of treatments. Following on from this, a profile of costs and consequences provides a representation that is readily comprehensible to guideline readers of any background. By explicitly identifying uncertainties, the presentation of the evidence accurately identifies strengths and weaknesses so that guideline development group members (and subsequently end users of the guideline report) can easily explore alternative values. The profile provides the starting point for a guideline development group to begin the process of valuing treatment alternatives and thereby producing recommendations. Sometimes further work may be needed using various forms of modelling to help a group to explore fully the meaning of the information before them.

As identified in chapter 3, the available evidence on which these presentations are based is not necessarily robust. Patient-oriented outcomes are reported particularly inconsistently in trials and it may be necessary to supplement meta-analytic clinical end-point findings with a narrative summary of quality-of-life findings where available. It may be necessary to access a wide range of sources to describe rare iatrogenic events, resource implications and unit cost data. For example, it was necessary to look at 3 years of coroners' findings in order to characterise rates of fatal poisoning associated with different antidepressants, since rates are too low to be captured within randomised controlled trials but toxicity is perceived to be a big concern.

There may be evidence that health-associated costs borne by patients and carers (e.g. travel and time to receive care, over-the-counter drugs, disability costs) and indirect costs of lost earnings differ significantly between alternative treatments. This should be considered relevant to a treatment decision at least in as much as it may undesirably influence adherence to treatment. There is the possibility that organisational alternatives may shift costs from the health service to individuals and the appropriateness of this may depend on the disease considered and contextual circumstances. Seldom are there adequate data to address costs borne by patients but where this is a concern these costs can be described as attributes of treatments.

While all of the 11 guideline development processes described support the rigorous identification of a range of evidence, they raise a series of generic issues: summarising study outcomes; time-frame issues; approaches to dealing with more complex disease areas; the development of profiles and models; and the role of health economists in guidelines.

Summarising study outcomes

Early on in the development of each of the guidelines there is a fundamental decision to be made about how to summarise the data from (usually) trials and whether or not there are

common outcomes across studies. If common outcomes are available then it may be possible to use quantitative techniques (meta-analysis or meta-regression) leading to summary relative and absolute estimates of benefit as discussed in chapter 3. If it is both sensible and possible to use quantitative methods, these allow an efficient presentation of available evidence. An issue arises as to whether available 'combinable' outcomes are those that the guideline development group identifies as important. For example, the trials incorporated in the post-myocardial infarction guideline routinely reported death rates but were more variable in their reporting of non-fatal cardiovascular outcomes. If available outcomes are not ideal then the group will have to decide whether or not the advantage of having a quantifiable measure of effect is worth the perceived disadvantage of working with an outcome measure that is not their first choice. However, it is only with a quantitative estimate of the effect of an intervention that it will be possible to quantify its cost-effectiveness.

If the evidence summary is to be qualitative (a narrative review of studies) the data can still be set out in ways that facilitate easy comparison between studies by using common descriptors (e.g. study design, study population, intervention, duration of intervention) using trial tables. However, under these circumstances it may not be possible to make estimates of cost-effectiveness unless the evidence summary is dominated by one trial with appropriate outcomes (this was the case for the diabetic foot ulcer guideline, where a large pragmatic trial was reported on the cost-effectiveness of prevention). In circumstances other than this, it is only possible to make cost-minimisation statements such those made in the asthma and angina guidelines: "as the treatments appear equivalent patients should use the cheapest preparation that they can tolerate and comply with". This situation is likely to arise when there is disparate evidence in terms of content or study design or insufficient skills or resources to use quantitative methods.

Time frame

The profile approach describes the attributes of alternative treatments in the time frame allowed by trial data. This can include the estimated net cost of treatments during this time frame and it may be possible to make presentations of (time-truncated) cost-effectiveness. Both costs and benefits are truncated, and it is possible that the cost-effectiveness ratio may be a reasonably robust estimator when ranking treatment alternatives and possibly less biased than one generated from extensive extrapolation assumptions, which require strong assumptions about the disease process. The advantage is that the economic presentation transparently reflects and summarises the available evidence profile. The doctor and patient must decide the likelihood that the benefits and costs described will apply in their context.

Although basing the profile on the trial time frame allows 'A' level treatment recommendations to be firmly based on the trial data, it leads to apparent anomalies. It is often the case (in chronic diseases) that clinicians wish to consider continuing treatment beyond the time window of the supporting evidence. This may lead to the situation where the treatment recommendation linked to the length of follow-up in trials is 'A' grade with the same recommendation being graded 'D' for treatment thereafter.[28] However, although this approach has been criticised,[117] it is only by such explicit linking of recommendations to the underlying evidence that both strengths and limitations become readily apparent.

Approach in more complex disease areas

There is no 'standard' sophistication of analysis that applies to every guideline: analysis will be influenced by several issues. Firstly the balance and nature of attributes found in the profile, secondly the breadth of the area under consideration, and additionally whether there is a need to consider multiple decisions, for example when sequencing treatment.

There will be occasions when a treatment provides clinical benefits with no (or infrequent) adverse effects and at little increased cost. This situation corresponds to Eddy's 'clear winner'.[15] An example would be aspirin as an anti-thrombotic in patients with ischaemic heart disease and in this guideline there was no need for extended consideration of costs. This was particularly so as the decision to prescribe or not was being considered in isolation from other management decisions for patients with ischaemic heart disease. However, as the situation becomes more complex the need for a fuller analysis generally becomes greater. A thorough analysis also pertains in a situation in which one is attempting to establish whether or not there is an evidence-based sequence to several treatments for one condition or several

treatments for several conditions. It is likely that a 'common playing field' will be needed to deal with features such as baseline risk in trials contributing evidence or dimensions of benefit reported inconsistently across treatments: such an analysis may well end up using modelling techniques.

It often proved possible from the profile of consequences and costs of treatments to reach clear treatment recommendations. This was often the case when a guideline addressed a particularly focussed or simple question, for example, should an ACE inhibitor be used in patients in primary care with confirmed heart failure. However, in other, similar, instances models were useful when important questions were left unanswered by the immediate evidence. For example, a guideline development group found a model helpful when exploring the effect on symptomatic ulceration and mortality of gastro-protective agents for osteoarthritis patients receiving NSAID therapy.

However, the usefulness of the profile may be severely tested when the range of comparisons is broad. This was exemplified in the guideline addressing available treatments for patients with prior myocardial infarction. In this instance there were significant differences between the different treatments in terms of the underlying severity of disease of the patients enrolled in trials. Consequently two treatments achieving the same relative reduction in mortality were achieving very different absolute improvements in survival during limited trial follow-up. The only path open to explore formally the concerns of the group was to model life-expectancy and the influence of baseline risk. It was necessary to set out the pros and cons of each approach carefully and spend time giving the group an overview of the modelling process and its assumptions. The profile and modelling approaches were received as mutually supportive presentations of the evidence, and raised the confidence of the group to reach firm recommendations. There was broad agreement that a profile may be more useful to a GP and the modelling exercise more helpful when thinking about a primary care group formulary. This emphasises the issue of the purpose of modelling of evidence within a guideline development process.

Profiles and modelling

Traditionally it is the province of health economics to model (combine, adjust, extrapolate, represent) intermediate clinical outcome data and data from other sources to explore the overall costs and consequences of treatment alternatives. In principle, it is possible to map clinical data onto generic quality-of-life scores, model the advancement of disease and produce cost/QALY estimates for each treatment decision. This was attempted in the development of the guideline for anti-coagulation treatment for patients with atrial fibrillation. However, this process contrasted with the other experiences in a number of ways. Firstly, there was no clear role for a multi-disciplinary guideline development group in deriving recommendations around the decision of whether or not to prescribe anti-coagulants – the 'right decision' was produced by the model. Secondly, the data were aggregated into a single metric, the constituent elements of which (and their associated uncertainty) were not transparent. Thirdly, the complexity of modelling a single decision was such that the viability of the method to deal with more complex clinical decisions, which have multiple interdependencies, has to be questioned.

The appropriate application of a decision analysis-driven guideline is currently unclear and a question for further research. However, this experience suggests that there are limits to how much additional information decision analysis-derived guidelines can convey (to either a guideline development group or subsequent users of the guideline) that would not be apparent in the presentation of the attributes of treatment. This belief stems from three realities. Firstly, the cost/QALY is not a useful metric to inform doctor–patient interactions (the *raison d'être* of a guideline). Where the whole aim of evidence-based guideline development is to make the process transparent, the production of an opaque summary statistic is less informative than an explicit profile. Secondly, a cost/QALY calculation inevitably uses treatment attribute data selectively and compounds the original data uncertainties with additional modelling uncertainties due to mapping, disease representation and (often) extrapolation beyond trial follow-up. There is no accepted way of reflecting these uncertainties and they are seldom adequately addressed in models, which leads to spurious precision in results. Thirdly, a cost/QALY calculation, by seeking to solve a big question – how does one treatment compare with any other treatment for any other disease? – may often be inefficient in terms of the objectives of a guideline. For example, knowing alternative antidepressants are similarly effective, tolerable and safe but have markedly different acquisition costs may be sufficient to

make clinically appropriate treatment recommendations.

However the issue is not a simple 'model/don't model' one. Meta-analysis is a kind of model that makes statistical and clinical assumptions when combining data. A profile of attributes used to help formulate treatment recommendations is one representation of reality. Although the assumptions may be less presumptuous than those found in many decision analyses, an implicit modelling process is being conducted when a guideline development group weighs up the attributes of treatments and formulates recommendations. The difference is one of emphasis: while health economists may seek to obtain clear, aggregated answers from a complex pattern of information, much of this pattern plays an important part in the clinical decision-making process. Often, simple modelling exercises have been necessary to enable the group to interpret the evidence before them. The touchstone for these exercises has been parsimony, to maximise the likelihood that subsequent readers of the full guideline report can follow, and if they wish, replicate or modify the process.

Skills utilised in guideline development and the role of the health economist

Although they have focused on the incorporation of economics within clinical guidelines, the case studies within this report have illustrated a more general development in the methods of guideline development. All of the guidelines described in chapters 4–7 were resourced with a set of core skills – where do the skills of the health economist sit within such an enterprise? The range of skills were those skills required to perform the tasks identified in chapter 3 and all of the guidelines had a health economist preparing data for the meetings. They also had a guideline methodologist who was responsible for managing the overall process of guideline development and leading the guideline development groups. All of the guidelines that used quantitative summary had a meta-analyst/health-services researcher as both a group member and a resource preparing data for the meetings. Both the meta-analysts and the health economist were supported by an information officer who performed the searches and organised the literature retrieval. Finally the whole process was underpinned by skilled secretarial support.

Such an appropriately skilled multi-disciplinary team supporting the process of guideline development has previously been identified as an important factor.[36] It has also been found that inexperienced guideline developers frequently underestimate the resources required for guideline development.[36] It is clear that different skills will be required more at some points than others. For example, the literature searching and retrieval will happen in the early stages of the process. The summarising of the effectiveness data will be a large tranche of work following on from this and the process of valuing the options will tend to happen towards the end. Early on in the guideline development group meetings, the health economist, meta-analyst and guideline methodologist have an educational role to fulfil, in addition to performing their 'technical' tasks, before a guideline development group can begin work. In practice this involves teaching groups sufficient of the topic areas to allow them to understand the interpretation of the evidence. However, the process relies on the interaction of all the players throughout the whole process. Therefore, if the aim is to produce clinical guidelines of the analytic complexity of those reported on page 20 (guideline on prevention and treatment of diabetic foot ulcers), then the resource and skill requirements are considerable.

Is it then necessary for every guideline development group to include a health economist? Williams[14] and Eddy[13] both argue that guidelines based on effectiveness issues and subsequently costed may differ substantially from, and be less efficient than, guidelines based on cost-effectiveness issues. Certainly, a health economist working independently would be deprived of the benefits of the interactions that occur within a guideline development group and may give personal values undue weight in any analysis. More fundamentally, the issue of the health economist's role exposes tensions between clinical care and health policy agendas and reflects a broader uncertainty about the target audience for economic analyses. Clinical guidelines are not primarily intended for health policy use: if resource implications and costs are to be considered as attributes of treatment in a valid manner and as contributors to valid recommendations, then this will have to occur as part of the guideline development process, not outside it.

For example, the profile of (non-cost) attributes of newer antidepressants left the guideline development group uncertain whether these drugs constituted a real advance: there were some pharmacological differences but in health

outcome terms the newer and older drugs were broadly equivalent. However, there were large differences in toxicity in (relatively infrequently occurring) overdose. If costs had not been considered, then in the face of equivalent effectiveness and different toxicity the decision would have been to recommend the use of newer drugs. However, when the group considered the cost implications of the newer drugs they were confident that it was not appropriate to recommend their first-line use. Similarly, the guideline development group considering the value of introducing donepezil hydrochloride for Alzheimer's disease concluded that the small clinical benefits of uncertain value to patients could not be countenanced in the light of the known and considerable costs. On the basis of effectiveness alone the group would have recommended the use of a new therapy with a small clinical effect. The impact of the important cost messages in both of these guidelines is likely to have been far less if they had been appended by a third party after the fact, rather than recommended by the group itself.

A health economist's participation in guideline development obtains a rich understanding of the implementation issues surrounding treatment that is rarely reflected in other forms of health technology assessment. An overview of guideline development issues for health economists about to embark on guideline development work is shown in *Box 16*.

Conclusions, implications and recommendations

No health economist would disagree that economic evaluation is about credibly valuing the alternative use of resources. This comes from a desire to minimise opportunity costs when making decisions about providing healthcare. The thinking behind the 'attribute profile' approach is data driven and seeks to open a participatory debate in a guideline development group about costs and consequences. The guideline process promotes a 'level playing field' where group members can participate as equals and in our experience they have responded very well to the process, demonstrating some profoundly 'economic' thinking in their treatment recommendations.

At its core, economics is about valuation, scarcity and choice. Decision-making occurs satisfactorily at the level of the individual when making choices about frequently purchased and competitively produced commodities. However, the nature of healthcare, as a publicly funded and provided commodity, is such that some form of social valuation is required. The guideline process leading to treatment recommendations might be considered to be an informed consensus valuation process. Clinical guidelines deliver 'ball-park' messages and do not address the allocation issue (i.e. how should resources be distributed between different diseases and treatments all competing for the same resources). Inadequate evidence both about the consequences of many treatments and society's valuation of healthcare are such that it is not currently possible to determine NHS provision on the basis of some grand efficiency scheme.

Implications

This report describes current optimum methods for developing clinical guidelines that include consideration of multiple dimensions of evidence (effectiveness, tolerability, harm, quality of life, resource implications, patient issues) and will be of relevance to those who commission, develop or use clinical guidelines. The described 'attribute profile' approach to judging whether the costs and consequences of treatments make reasonable sense appears to be the most robust and socially defensible method at this time.

The main implication from this work is that these methods should form the current minimum expected of guideline developers. It is important that the methods described are attempted and developed by other guideline methodologists and health economists and the debate about the social valuation of healthcare is expanded.

Recommendations for further research

In working on the case studies a range of unanswered questions have been identified, some of which are directly related to the notion of considering costs within guidelines and some of which relate to clinical guideline development more generally.

- Given that the taxonomies for level of evidence and strength of recommendation are largely derived around issues of effectiveness and tolerability, what are the implications of considering a wider range of treatment attributes (costs, patients' experiences) for the categorisation of evidence and recommendations?
- What is the relationship between the incorporation of costs into a guideline (as one of a range of treatment attributes) and the cost impact of a guideline (that may or may not consider the costs of active implementation). What are the

BOX 16 An overview of guideline development for health economists

- It is important to be clear about the process and objectives of guideline work, the conduct of group meetings and the role of each of the group members
- Objective (if probabilistic) attributes of treatment decisions include effectiveness, side-effects, compliance, safety, quality of life, health-service delivery issues, resource use and costs. The outcomes of a guideline process are graded treatment recommendations, which may reflect some or all of these attributes
- The health economist, together with other group facilitators, has a responsibility of providing a rigorous exploration of treatment attributes with the available evidence. A general understanding of other disciplines (statistics, epidemiology and health-services research) is essential alongside training in economic evaluation methodology
- Simple and transparent presentations, which permit exploration with different values, are most likely to be helpful to the guideline development group and subsequent users of the guideline
- Each attribute of treatment is assessed in turn on its own merits including bounds of uncertainty. Over-precision should be challenged and all uncertainties that are appropriate to the data or expressed in the group should be explored
- Careful presentation and full discussion in the guidelines group is essential for an understanding of the attributes of treatment to evolve into a view of overall clinical importance. Data, although rigorously analysed, are being used to put a treatment in a broad 'ball-park' with respect to its various attributes (e.g. safe, acceptable, effective, deliverable, cost neutral)
- The importance attached to each attribute of treatment remains the responsibility of the guideline development group as a whole and recommendations must be transparent and credible to the target audience. A summary table of attributes, presenting summary cost-effectiveness estimates when appropriate, may facilitate the process of aggregating up to an overall recommendation. Summary ratios need careful explanation and interpretation
- A systematic review of previously conducted economic analyses relating to a treatment may provide useful background to the health economist but may have limited (or no) direct use in the guidelines process. If rigorously conducted, the guideline process is likely to produce an understanding and evaluation of treatment inadequately reflected in any one published analysis
- The scope for conducting traditional cost-effectiveness models may be limited or unhelpful in some therapeutic areas, especially where the various attributes of treatments contain conflicting messages. Where modelling is appropriate, clinicians appear more responsive to simple and transparent models, than complex 'black box' methods requiring greater assumptions and extrapolation
- Grading of recommendations of cost-effectiveness is in its infancy. However, the grade of a recommendation should reflect not only precision and susceptibility to bias but also generalisability and clinical relevance
- As with clinical effectiveness, in a recommendation about the cost-effectiveness of treatments it is acceptable to say that 'we just don't know yet' and that further research is required

optimum methods of using cost data in guideline development and of assessing the cost impact of a guideline? Should they be conducted within the same, or separate, processes?

- All the methods described within this report are derived from case studies. A number of the uncertainties identified about the optimum methods could be resolved by direct comparison of methods within the development of clinical guidelines. This could address both the impact of methods within the development of guidelines and the impact on subsequent value to both clinicians and patients.
- It is unclear what role there is for decision analysis in the development of clinical guidelines. Further work could usefully clarify: the ability of decision analysis to cope with the scale and scope of the multiple decisions required of a disease-based clinical guideline; the impact of the lack of transparency of the process on subsequent use; the usefulness of collecting patient utility scores in routine clinical decision-making.
- In what circumstances is it necessary to use formal consensus methods within a guideline development process?
- What are the optimum methods of incorporating consumers' views within the development of clinical guidelines?

Acknowledgements

This project had input from a number of people. We are particularly grateful to Professor Mike Drummond, Professor Jeremy Grimshaw, Professor Allen Hutchinson and Professor Nick Freemantle for their input into the content of the guideline case studies and sections of this report. We would also like to thank Dr Paul Shekelle and Professor Steve Woolf for their thoughtful comments and correspondence on guidelines over the time that these guideline case studies were conducted.

We wish to express our thanks to all of the health-service professionals, specialist resources, support staff and external referees who gave their time and energy to the guideline processes described in this report.

We are grateful to the DARTS Study Group for their permission to include the case study of the evaluative guidelines for anticoagulation to prevent stroke in patients with atrial fibrillation. (DARTS Study Group: from the Department of Epidemiology and Public Health, School of Health Sciences, University of Newcastle upon Tyne – Richard Thomson, Professor of Epidemiology and Public Health; Dr David Parkin, Senior Lecturer in Health Economics; Dr Mark Sudlow, Medical Research Council Special Health Services Research Training Fellow; Angela Robinson, Research Associate; from the Centre for Health Services Research, School of Health Sciences, University of Newcastle upon Tyne – Martin Eccles, Professor of Clinical Effectiveness).

The development of the guidelines described in this report were variously funded by the Department of Health for England and Wales and the Northern and Yorkshire R&D Programme.

The Centre for Health Services Research, University of Newcastle Upon Tyne and the Centre for Health Economics, University of York are both members of the Medical Research Council Health Services Research Collaboration.

We are indebted to the referees for their perseverance in reading this report and the quality of their comments.

The views expressed in this report are those of the authors, who are also responsible for any errors.

References

1. Institute of Medicine. Guidelines for clinical practice: from development to use. Washington: National Academy Press; 1992.
2. Woolf SH, Grol R, Hutchinson A, Eccles M, Grimshaw J. Clinical guidelines: the potential benefits, limitations and harms of clinical guidelines. *BMJ* 1999;**318**:527–30.
3. Grol R, Eccles M, Maisonneuve H, Woolf S. Developing clinical practice guidelines: the European experience. *Dis Manage Health Outcomes* 1998;**4**(5):255–66.
4. Woolf S, Grol R, Hutchinson A, Eccles M, Grimshaw J. An international overview. In: Eccles M, Grimshaw J, editors. Clinical guidelines: from conception to use. Oxford: Radcliffe Medical Press; 2000. p. 31–48.
5. NHS Executive. Clinical guidelines: using clinical guidelines to improve patient care within the NHS. London: HMSO; 1996.
6. Department of Health. The new NHS: modern, dependable. London: Department of Health; 1997.
7. NHS Executive. A first class service: quality in the new NHS. London: HMSO; 1998.
8. Grimshaw J, Russell I. Achieving health gain through clinical guidelines I: developing scientifically valid guidelines. *Qual Health Care* 1993;**2**:243–8.
9. Eccles MP, Clapp Z, Grimshaw J, Adams PC, Higgins B, Purves I, *et al.* Developing valid guidelines: methodological and procedural issues from the North of England Evidence Based Guideline Development Project. *Qual Health Care* 1996;**5**:44–50.
10. Grimshaw J, Eccles M, Russell I. Developing clinically valid practice guidelines. *J Eval Clin Pract* 1995;**1**(1):37–48.
11. Egger M, Davey Smith G, Altman DG. Systematic reviews in health care. Meta-analysis in context. London: BMJ Publishing Group; 2001.
12. Eccles M, Higgins B, Adams P. Developing evidence-based guidelines: experiences from the North of England Evidence-Based Guideline Development Group. In: Hutchinson A, Baker R, editors. Making use of guidelines in clinical practice. Oxford: Radcliffe Medical Press; 1999. p. 29–46.
13. Eddy DM. A manual for assessing health practices and designing practice policies: the explicit approach. Philadelphia: American College of Physicians; 1992.
14. Williams A. How should information on cost effectiveness influence clinical practice? In: Delamothe T, editor. Outcomes into clinical practice. London: BMJ Publishing Group; 1995. p. 99–107.
15. Eddy DM. Doctors, economics and clinical practice guidelines: can they be brought together? London: Office of Health Economics; 1999.
16. Clinical Guidelines Working Group. The development and implementation of clinical guidelines. Report from General Practice 26. Exeter: Royal College of General Practitioners; 1995.
17. Royal College of Psychiatrists. Clinical practice guidelines and their development. London: Royal College of Psychiatrists; 1994. Council Report CR34.
18. Eccles MP, Clapp Z, Grimshaw JM, Adams PC, Higgins B, Purves I, *et al.* North of England Evidence Based Guidelines Development Project: methods of guideline development. *BMJ* 1996;**312**:760–1.
19. Eccles M, Freemantle N, Mason J. North of England Evidence Based Guidelines Development Project: methods of developing guidelines for efficient drug use in primary care. *BMJ* 1998;**316**:1232–5.
20. Eccles M, Grimshaw J. Clinical guidelines: from conception to use. Oxford: Radcliffe Medical Press; 2000.
21. Woolf SH. An organized analytic framework for practice guideline development: using the analytic logic as a guide for reviewing evidence, developing recommendations, and explaining the ration. In: McCormick KA, Moore SR, Siegel RA, editors. Methodology perspectives. Washington DC: Agency for Health Care Policy and Research; 1994. p. 105–13.
22. Eccles M, Freemantle N, Mason JM. North of England Evidence Based Guideline Development Project: summary version of guidelines for the choice of antidepressants for depression in primary care. *Fam Pract* 1999;**16**:103–11.
23. Eccles M, Freemantle N, Mason J. North of England Evidence Based Guideline Development Project: summary guideline for non-steroidal anti-inflammatory drugs versus basic analgesia in treating the pain of degenerative arthritis. *BMJ* 1998;**317**:526–30.

24. North of England Evidence Based Guideline Development Project. Evidence based clinical practice guideline: non-steroidal anti-inflammatory drugs (NSAIDs) versus basic analgesia in the treatment of pain believed to be due to degenerative arthritis. Report 86. Newcastle upon Tyne: Centre for Health Services Research; 1998.

25. Eccles M, Freemantle N, Mason J, North of England ACE-inhibitor Guideline Development Group. North of England Evidence Based Development Project: guideline for angiotensin converting enzyme inhibitors in the primary care management of adults with symptomatic heart failure. *BMJ* 1998;**316**:1369–75.

26. North of England Evidence Based Guideline Development Project. Evidence based clinical practice guideline: ACE inhibitors in the primary care management of adults with symptomatic heart failure. Report 90. Newcastle upon Tyne: Centre for Health Services Research; 1998.

27. Eccles M, Freemantle N, Mason J, North of England Aspirin Guideline Development Group. North of England Evidence Based Guidelines Development Project: evidence based guideline for the use of aspirin for the secondary prophylaxis of vascular disease in primary care. *BMJ* 1998;**316**:1303–9.

28. North of England Evidence Based Guideline Development Project. Evidence based clinical practice guideline: the use of aspirin for the secondary prophylaxis of vascular disease in primary care. Report 89. Newcastle upon Tyne: Centre for Health Services Research; 1998.

29. Eccles M, North of England Evidence Based Guideline Development Project. North of England Evidence Based Guidelines Development Project: summary version of evidence based guideline for the primary care management of dementia. *BMJ* 1998;**317**:802–8.

30. North of England Evidence Based Guideline Development Project. Evidence based clinical practice guideline: the primary care management of dementia. Report 92. Newcastle upon Tyne: Centre for Health Services Research; 1998.

31. Kahan JP, Park RE, Bernstein SJ, Hilborne LH, Parker L, *et al.* Variations by specialty in physician ratings of the appropriateness and necessity of indications for procedures. *Med Care* 1996;**34**(6):512–23.

32. Coulter I, Adams A, Shekelle P. Impact of varying panel membership on ratings of appropriateness in consensus panels – a comparison of a multi- and single disciplinary panel. *Health Serv Res* 1995;**30**(4):577–91.

33. Leape LL, Park RE, Kahan JP, Brook RH. Group judgments of appropriateness: the effect of panel composition. *Qual Assur Health Care* 1992; **4**(2):151–9.

34. Kosecoff JH, Kanouse DE, Rogers WH, McCloskey L, Winslow CM, Brook RH. Effects of the National Institutes of Health Consensus Development Program on physician practice. *JAMA* 1987;**258**(19):2708–13.

35. Shekelle PG, Schriger DL. Evaluating the use of the appropriateness method in the Agency for Health Care Policy and Research Clinical Practice Guideline Development process. *Health Serv Res* 1996;**31**(4):453–68.

36. Shekelle PG, Woolf SH, Eccles M, Grimshaw J. Developing guidelines. *BMJ* 1999;**318**:593–6.

37. Murphy MK, Black NA, Lamping DL, McKee CM, Sanderson CFB, Askham J, *et al.* Consensus development methods, and their use in clinical guideline development. *Health Technol Assess* 1998;**2**(3).

38. Standing Advisory Group on Consumers in NHS Research. Involvement works. Winchester: NHS Executive; 1999.

39. Van Wersch A, Eccles M. Patient involvement in evidence-based health in relation to clinical guidelines. In: Gabbay M, editor. The evidence-based primary care handbook. London: Royal Society of Medicine Press Ltd; 1999. p. 91–103.

40. Van Wersch A, Eccles M. Involvement of consumers in the development of evidence based clinical guidelines: practical experiences from the North of England Evidence Based Guideline Development Programme. *Qual Health Care* 2001;**10**:10–16.

41. Eccles M, Mason J, Freemantle N. Developing valid cost effectiveness guidelines: a methodological report from the North of England evidence based guideline development project. *Qual Health Care* 2000;**9**(2):127–32.

42. Cook TD, Campbell DT. Quasi-experimentation: design and analysis issues for field settings. Chicago: Rand McNally; 1979.

43. US Department of Health and Human Services, Public Health Service, Agency for Health Care and Policy Research. Acute pain management: operative or medical procedures and trauma. Rockville, MD: Agency for Health Care and Policy Research Publications; 1992.

44. Begg CB, Berlin JA. Publication bias: a problem in interpreting medical data. *J R Stat Soc* 1988; **151**:419–63.

45. Easterbrook PJ, Berlin JA, Gopalan R, Matthews DR. Publication bias in clinical research. *Lancet* 1991;**337**(13 April):867–72.

46. Freemantle N, Eccles M, Mason J. Deriving treatment recommendations from evidence within randomized trials: the role and limitation of meta-analysis. *Int J Technol Assess Health Care* 2000;**15**(2):304–15.

47. SOLVD Investigators. Effect of enalapril on survival in patients with reduced left ventricular ejection fractions and congestive heart failure. *N Engl J Med* 1991;**325**:293–302.

48. Hedges LV, Olkin I. Statistical methods for meta analysis. London: Academic Press; 1985.

49. Freemantle N, Drummond MF. Should clinical trials with concurrent economic analyses be blinded? *JAMA* 1997;**277**:63–4.

50. Smith J, Channer KS. Increasing prescription of drugs for secondary prevention after myocardial infarction. *BMJ* 1995;**311**:917–18.

51. Mooney GH, Drummond MF. Essentials of health economics: part 1 (continued) – what is economics? *BMJ* 1982;**285**:1024–5.

52. Mason J, Eccles M, Freemantle N, Drummond M. A framework for incorporating cost-effectiveness in evidence based clinical practice guidelines. *Health Policy* 1999;**47**:37–52.

53. Freemantle N, Mason JM. Publication bias in clinical trials and economic analyses. *Pharmacoeconomics* 1997;**12**:10–16.

54. Mason JM. The generalisability of pharmaco-economic studies. *Pharmacoeconomics* 1997;**11**:503–14.

55. North of England Evidence Based Guideline Development Project. The primary care management of stable angina. Report 98. Newcastle upon Tyne: Centre for Health Services Research, Evidence Based Clinical Practice Guideline; 1999.

56. North of England Stable Angina Guideline Development Group. North of England evidence based guidelines development project: summary version of evidence based guideline for the primary care management of stable angina. *BMJ* 1996;**312**:827–32.

57. North of England Evidence Based Guideline Development Project. The primary care management of asthma in adults. Report 97. Newcastle upon Tyne: Centre for Health Services Research, Evidence Based Clinical Practice Guideline; 1999.

58. North of England Asthma Guideline Development Group. North of England evidence based guidelines development project: summary version of evidence based guideline for the primary care management of asthma in adults. *BMJ* 1996;**312**:762–6.

59. Hutchinson A, McIntosh A, Feder G, Home PD, Mason J, O'Keeffe C, *et al.* Clinical guidelines and evidence review for Type 2 diabetes: prevention and management of foot problems. London: Royal College of General Practitioners; 2000.

60. Mason J, O'Keeffe C, McIntosh A, Hutchinson A, Booth A, Young RJ. A systematic review of foot ulcer in patients with Type 2 diabetes mellitus. 1: prevention. *Diabetic Med* 1999;**16**:801–12.

61. Mason J, O'Keeffe C, Hutchinson A, McIntosh A, Young R, Booth A. A systematic review of foot ulcer in patients with Type 2 mellitus. II: treatment. *Diabetic Med* 1999;**16**:889–909.

62. McCabe CJ, Stevenson RC, Dolan AM. Evaluation of a diabetic foot screening and protection programme. *Diabetic Med* 1998;**15**:80–4.

63. Kumar S, Fernando DJS, Veves A, Knowles EA, Young MJ, Boulton AJM. Semmes-Weinstein monofilaments: a simple, effective and inexpensive screening device for identifying diabetic patients at risk of foot ulceration. *Diabetes Res Clin Pract* 1991;**13**:63–8.

64. Joint Formulary Committee. British National Formulary. 38th ed. London: British Medical Association and Royal Pharmaceutical Society of Great Britain; 1999.

65. McMurray J, Hart W. The economic impact of heart failure on the UK National Health Service. *Eur Heart J* 1993;**14**(Suppl):133.

66. Kleber FX. Socio-economic aspects of ACE inhibition in the secondary prevention in cardiovascular diseases. *Am J Hypertension* 1994;**7**:1125–65.

67. Paul SD, Kuntz KM, Eagle KA, Weinstein MC. Costs and effectiveness of angiotensin converting enzyme inhibition in patients with congestive heart failure. *Arch Intern Med* 1994;**154**:1143–9.

68. Van Hout BA, Wielink G, Bonsel GJ, Rutten FHF. Effects of ACE inhibitors on heart failure in the Netherlands: a pharmacoeconomic model. *Pharmacoeconomics* 1993;**3**:387–97.

69. Hart W, Rhodes G, McMurray J. The cost effectiveness of enalapril in the treatment of chronic heart failure. *Br J Med Econ* 1993;**6**:91–8.

70. Joint Formulary Committee. British National Formulary. 32nd ed. London: British Medical Association and Royal Pharmaceutical Society of Great Britain; 1996.

71. Drummond MF, McGuire A, Fletcher A. Economic evaluation of drug therapy for hyper-cholesterolaemia in the United Kingdom. Discussion paper 104. University of York: Centre for Health Economics; 1993.

72. Freemantle N, House A, Song F, Mason J, Sheldon TA. Prescribing selective serotonin reuptake inhibitors as strategy for prevention of suicide. *BMJ* 1994;**309**:249–53.

73. Netten A, Dennett J. Unit costs of health and social care. University of Canterbury: Personal Social Services Research Unit (PPSRU); 1997.

74. The Health Service Financial Database. London: Chartered Institute of Public Finance and Accountancy; 1997.

75. Antiplatelet Trialists' Collaboration. Collaborative overview of randomised trials of antiplatelet therapy –I: prevention of death, myocardial infarction, and stroke by prolonged antiplatelet therapy in various categories of patients. *BMJ* 1994;**308**:81–106.

76. Roderick PJ, Wilkes HC, Meade TW. The gastrointestinal toxicity of aspirin: an overview of randomised controlled trials. *Br J Clin Pharmacol* 1993;**35**:219–26.

77. Aspirin Myocardial Infarction Study (AMIS) Research Group. AMIS: a randomized controlled trial of aspirin in persons recovered from myocardial infarction. *JAMA* 1980;**243**:661–9.

78. Aspirin Myocardial Infarction Study (AMIS) Research Group. AMIS: the aspirin myocardial infarction study: final results. *Circulation* 1980;**62**(Suppl V):79–84.

79. UK-TIA Study Group. United Kingdom Transient Ischaemic Attack (UK-TIA) aspirin trial: interim results. *BMJ* 1988;**296**:316–20.

80. UK-TIA Study Group. United Kingdom Transient Ischaemic Attack (UK-TIA) aspirin trial: final results. *J Neurol Neurosurg Psychiatry* 1991; **54**:1044–54.

81. SALT Collaborative Group. Swedish Aspirin Low-dose Trial (SALT) of 75 mg aspirin as secondary prophylaxis after cerebrovascular ischaemic events. *Lancet* 1991;**338**:1345–9.

82. Juul-Möller S, Edvardsson N, Jahnmatz B, Rosen A, Sorensen S, Ömblus R, *et al.* Double blind trial of aspirin in primary prevention of myocardial infarction in patients with stable chronic angina pectoris. The Swedish Angina Pectoris Aspirin Trial (SAPAT) Group. *Lancet* 1992;**340**:1421–5.

83. North of England Evidence Based Guideline Development Project. Evidence based clinical practice guideline: the choice of anti-depressants for depression in primary care. Report 91. Newcastle upon Tyne: Centre for Health Services Research; 1998.

84. National Health Service in Scotland. Scottish Health Service costs. Edinburgh: National Health Service in Scotland; 1996.

85. Jönsson B, Bebbington P. What price depression? The cost of depression and the cost effectiveness of pharmacological treatment. *Br J Psychiatry* 1994;**164**:665–73.

86. Freemantle N, House A, Mason J, Song F, Sheldon T. Economics of treatment of depression [Letter]. *Br J Psychiatry* 1995;**166**:397–405.

87. Jönsson B, Bebbington P. Economics of treatment of depression [Author's reply]. *Br J Psychiatry* 1995;**166**:398–9.

88. Woods SW, Rizzo JA. Cost-effectiveness of anti-depressant treatment reassessed. *Br J Psychiatry* 1997;**170**:257–63.

89. Simon GE, VonKorff M, Heilingenstein JH, Revicki DA, Grothaus L, Katon W, *et al.* Initial antidepressant choice in primary care: effectiveness and cost of fluoxetine vs tricyclic antidepressants. *JAMA* 1996;**275**:1897–902.

90. Department of Health. Health and personal social services statistics in England. A publication of the Government Statistical Service. London: Department of Health; 1996.

91. Henry D, Lim LL-Y, Garcia Rodriguez LA, Perez Gutthann S, Carson JL, Griffin M, *et al.* Variability in risk of gastrointestinal complications with individual non-steroidal anti-inflammatory drugs: results of a collaborative meta analysis. *BMJ* 1996;**312**:1563–6.

92. Stucki G, Johannesson M, Liang MH. Is misoprostol cost-effective in the prevention of non-steroidal anti-inflammatory drug-induced gastropathy in patients with chronic arthritis. *Arch Intern Med* 1994;**154**:2020–5.

93. Hillman AL, Bloom BS. Economic effects of prophylactic use of misoprostol to prevent gastric ulcer in patients taking non-steroidal anti-inflammatory drugs. *Arch Intern Med* 1989;**149**:2061–5.

94. Knill-Jones R, Drummond M, Kohli H, Davies L. Economic evaluation of gastric ulcer prophylaxis in patients with arthritis receiving non-steroidal anti-inflammatory drugs. *Postgrad Med J* 1990;**66**:639–46.

95. Jönsson B, Haglund U. Cost-effectiveness of misoprostol in Sweden. *Int J Tech Assess Health Care* 1992;**8**:234–44.

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

97. Edelson JT, Tosteson ANA, Sax P. Cost-effectiveness of misoprostol for prophylaxis against non-steroidal anti-inflammatory drug-induced gastrointestinal tract bleeding. *JAMA* 1990;**264**:41–7.

98. Drummond MF, Bloom BS, Carrin G, Hillman AL, Hutchings HC, Knill-Jones RP, *et al.* Issues in the cross-national assessment of health technology. *Int J Tech Assess Health Care* 1992;**8**:671–82.

99. Graham DY, Agrawal NM, Roth SH. Prevention of NSAID-induced gastric ulcer with misoprostol: multicentre, double-blind, placebo-controlled trial. *Lancet* 1998;**2**:1277–80.

100. Silverstein FE, Graham DY, Senior JR, Davies HW, Struthers BJ, Bittman RM, *et al.* Misoprostol reduces serious gastrointestinal complications in patients with rheumatoid arthritis receiving nonsteroidal anti-inflammatory drugs: a randomized, double-blind, placebo-controlled trial. *Ann Intern Med* 1995;**123**:241–9.

101. Rogers SL, Friedhoff LT, the Donepezil Study Group. The efficacy and safety of donepezil in patients with Alzheimer's disease: results of a US multicentre randomised double blind, placebo-controlled trial. *Dementia* 1996;**7**:293–303.

102. McCormick A, Fleming D, Charlton J. Morbidity statistics from general practice. Fourth national study 1991–1992. London: OPCS, HMSO; 1995.

103. Chew R. OHE Compendium of Health Statistics. 9th ed. London: Office of Health Economics; 1995.

104. Department of Health and Office of Population Censuses and Surveys. The government's expenditure plans 1994–5 to 1996–7. London: HMSO; 1994.

105. Gray A, Fenn P. Alzheimer's disease: the burden of the illness in England. *Health Trends* 1993;**25**:31–37.

106. Geddes J, Freemantle N, Harrison P, Bebbington P. Atypical antipsychotics in the treatment of schizophrenia: systematic overview and meta-regression analysis. *BMJ* 2000;**321**:1371–6.

107. Rosenheck R, Cramer J, Xu W, Thomas J, Henderson W, Frisman L, *et al.* A comparison of clozapine and haloperidol in hospitalized patients with refractory schizophrenia. *N Engl J Med* 1997;**337**:809–15.

108. Meredith C, Mahmoud RA, Englehart LM, Ramirez L, Risperidone Outcome Study of Effectiveness (ROSE) Group. Clinical and quality of life superiority of risperidone over conventional anti-psychotics under usual care conditions: a prospective randomized trial in schizophrenia and schizoaffective disorder. Annual Meeting of the American Psychiatric Association; 1998 May 30–June 4; Toronto, Canada.

109. Tollefson GD, Beasley CM, Tamura RN, Tran PV, Potvin JH. Blind, controlled, long-term study of the comparative incidence of treatment emergent tardive dyskinesia with olanzapine or haloperidol. *Am J Psychiatry* 1997;**154**:1248–54.

110. The Development and Evaluation Committee. The use of olanzapine as a first and second choice treatment in schizophrenia. Report No 84. Bristol: NHS R&D; 1998.

111. The Heart Outcomes Prevention Evaluation (HOPE) Study Investigators. Effects of an angiotensin-converting-enzyme inhibitor, ramipril, on cardiovascular events in high-risk patients. *N Engl J Med* 2000;**342**:145–53.

112. Long-Term Intervention with Pravastatin in Ischaemic Disease (LIPID) Study Group. Prevention of cardiovascular events and death with pravastatin in patients with coronary heart disease and a broad range of initial cholesterol levels. *N Engl J Med* 1998;**339**:1349–57.

113. β-Blocker Heart Attack Trial (ß-HAT) Research Group. A randomized trial of propranolol in patients with acute MI. 1. Mortality results. *JAMA* 1982;**247**:1707–14.

114. Pitt B, Zannad F, Remme WJ, Cody R, Castaigne A, Perez A, *et al.* The effect of spironolactone on morbidity and mortality in patients with severe heart failure. *N Engl J Med* 1999;**341**:709–17.

115. Thomson R, Parkin D, Eccles M, Sudlow M, Robinson A. Evaluative development of clinical guidelines on the use of treatments to prevent stroke amongst patients with atrial fibrillation for use in the Northern & Yorkshire Region. Newcastle upon Tyne: School of Health Sciences, University of Newcastle; 1998.

116. Thomson R, Parkin D, Eccles M, Sudlow M, Robinson A. Decision analysis and guidelines for anticoagulant therapy to prevent stroke in patients with atrial fibrillation. *Lancet* 2000;**55**:956–62.

117. Chant H, Fearn S, McCollum C. Aspirin prophylaxis for vascular disease. Carotid endarterectomy should have been mentioned [Letter]. *BMJ* 1998; **317**:1318–19.

Methodology Group

Members

Methodology Programme Director
Professor Richard Lilford
Director of Research and Development
NHS Executive – West Midlands, Birmingham

Chair
Professor Martin Buxton
Director, Health Economics Research Group
Brunel University, Uxbridge

Professor Douglas Altman
Professor of Statistics in Medicine
University of Oxford

Dr David Armstrong
Reader in Sociology as Applied to Medicine
King's College, London

Professor Nicholas Black
Professor of Health Services Research
London School of Hygiene & Tropical Medicine

Professor Ann Bowling
Professor of Health Services Research
University College London Medical School

Professor David Chadwick
Professor of Neurology
The Walton Centre for Neurology & Neurosurgery
Liverpool

Dr Mike Clarke
Associate Director (Research)
UK Cochrane Centre, Oxford

Professor Paul Dieppe
Director, MRC Health Services Research Centre
University of Bristol

Professor Michael Drummond
Director, Centre for Health Economics
University of York

Dr Vikki Entwistle
Senior Research Fellow, Health Services Research Unit
University of Aberdeen

Professor Ewan B Ferlie
Professor of Public Services Management
Imperial College, London

Professor Ray Fitzpatrick
Professor of Public Health & Primary Care
University of Oxford

Dr Naomi Fulop
Deputy Director, Service Delivery & Organisation Programme
London School of Hygiene & Tropical Medicine

Mrs Jenny Griffin
Head, Policy Research Programme
Department of Health
London

Professor Jeremy Grimshaw
Programme Director
Health Services Research Unit
University of Aberdeen

Professor Stephen Harrison
Professor of Social Policy
University of Manchester

Mr John Henderson
Economic Advisor
Department of Health, London

Professor Theresa Marteau
Director, Psychology & Genetics Research Group
Guy's, King's & St Thomas's School of Medicine, London

Dr Henry McQuay
Clinical Reader in Pain Relief
University of Oxford

Dr Nick Payne
Consultant Senior Lecturer in Public Health Medicine
ScHARR
University of Sheffield

Professor Joy Townsend
Director, Centre for Research in Primary & Community Care
University of Hertfordshire

Professor Kent Woods
Director, NHS HTA Programme, & Professor of Therapeutics
University of Leicester

continued

HTA Commissioning Board

Members

Programme Director
Professor Kent Woods
Director, NHS HTA Programme, &
Professor of Therapeutics
University of Leicester

Chair
Professor Shah Ebrahim
Professor of Epidemiology of Ageing
University of Bristol

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

Professor Douglas Altman
Director, ICRF Medical Statistics Group
University of Oxford

Professor John Bond
Director, Centre for Health Services Research
University of Newcastle-upon-Tyne

Ms Christine Clark
Freelance Medical Writer
Bury, Lancs

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

Dr Andrew Farmer
General Practitioner &
NHS R&D Clinical Scientist
Institute of Health Sciences
University of Oxford

Professor Adrian Grant
Director, Health Services Research Unit
University of Aberdeen

Dr Alastair Gray
Director, Health Economics Research Centre
Institute of Health Sciences
University of Oxford

Professor Mark Haggard
Director, MRC Institute of Hearing Research
University of Nottingham

Professor Jenny Hewison
Senior Lecturer
School of Psychology
University of Leeds

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

Dr Donna Lamping
Head, Health Services Research Unit
London School of Hygiene & Tropical Medicine

Professor David Neal
Professor of Surgery
University of Newcastle-upon-Tyne

Professor Gillian Parker
Nuffield Professor of Community Care
University of Leicester

Dr Tim Peters
Reader in Medical Statistics
University of Bristol

Professor Martin Severs
Professor in Elderly Health Care
University of Portsmouth

Dr Sarah Stewart-Brown
Director, Health Services Research Unit
University of Oxford

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

Dr Gillian Vivian
Consultant in Nuclear Medicine & Radiology
Royal Cornwall Hospitals Trust
Truro

Professor Graham Watt
Department of General Practice
University of Glasgow

Dr Jeremy Wyatt
Senior Fellow
Health Knowledge Management Centre
University College London

Current and past membership details of all HTA 'committees' are available from the HTA website (see inside front cover for details)

Feedback

The HTA programme and the authors would like to know your views about this report.

The Correspondence Page on the HTA website (http://www.ncchta.org) is a convenient way to publish your comments. If you prefer, you can send your comments to the address below, telling us whether you would like us to transfer them to the website.

We look forward to hearing from you.

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) 23 8059 5639 Email: hta@soton.ac.uk
http://www.ncchta.org

ISSN 1366-5278