The Journal of Mental Health Policy and Economics
J. Mental Health Policy Econ. **2**, 13–19 (1999)

Incorporating Economic Analysis in Evidence-Based Guidelines for Mental Health: the Profile Approach

James Mason*[1], Martin Eccles[2], Nick Freemantle[1] and Michael Drummond[1]

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

Abstract

Background: Many western health systems are currently developing the role of clinical guidelines to promote effective and efficient health care. However, introducing economic data into guideline methodology designed to assess the effectiveness of interventions raises some methodological issues. These include providing valid and generalizable cost estimates, the weight placed upon cost 'evidence' and presenting cost-effectiveness information in a way that is helpful to clinicians.
Aim of the Study: To explore a framework for including economic concepts in the development of a series of primary care guidelines, two of which address mental health conditions.
Methods: A profile approach, setting out best available evidence about the attributes of treatment choices (effectiveness, tolerability, safety, health service delivery, quality of life, resource use and cost), was used to help clinicians to derive treatment recommendations in a manner consistent with both the clinical decision-making process and social objectives.
Results: Clinicians involved in guideline development responded well to the process. Although there was often considerable debate about the meaning and importance of different aspects of evidence about treatment, in none of the guideline groups was there failure to agree treatment recommendations.
Discussion: The profile approach may be particularly useful in the field of mental health where disease processes may often feature very disparate effects, over long periods of time and impacting upon a broad circle of relatives, carers and agencies in addition to the patients themselves.
Conclusion: A method has been applied in a series of primary care guidelines, which appears to enable clinicians to consider the issue of resource use alongside the various clinical attributes associated with treatment decisions. The basis of this work is the belief that guidance presenting physical measures describing effectiveness, adverse events, safety, compliance and quality of life, alongside resource consequences, is most likely to appropriately inform doctor–patient interactions.
Implications for Health Care Provision and Use: This research may provide a useful platform for other groups considering how to introduce cost-effectiveness concepts into guideline development groups. Whether guidelines change clinical behaviour remains a research question, and the subject of forthcoming trials.
Implications for Health policy Formulation: It is important that government agencies realize that guideline development is a health policy tool with prescribed methods to produce valid guidelines. Attempts to tamper with the methodology for cost-containment purposes or other political reasons are likely to discredit a useful mechanism for improving the scientific basis of health care provision.
Implications for Further Research: There are a number of limitations to completed work: for example it has a primary care focus and addresses fairly narrowly defined conditions. Work is ongoing to extend the scope to broader disease areas and to secondary care. Copyright © 1999 John Wiley & Sons, Ltd.

Received 10 October 1998; accepted 22 February 1999.

Introduction

Guideline methodology has evolved rapidly over the last decade from simple consensus or opinion pieces into a highly structured approach for assessing and summarizing evidence and deriving appropriate treatment recommendations.[1,2] The incorporation of health economics within guidelines has been argued for from a number of sources[3–6] although no criteria have hitherto been developed to judge whether this has been achieved appropriately.

The new policy direction for the English health service puts great emphasis on the development of the managed care of diseases, audit and self-regulation.[7] A substantial input to these processes will be achieved through the implementation of clinical and cost-effectiveness guidelines. A new National Institute for Clinical Excellence (NICE) is being established to help present this information coherently to the NHS.[8] These changes put renewed emphasis on the need to find an appropriate structure for cost-conscious guidelines.

Many health care professionals have a limited knowledge and innate mistrust of health economics and economic modelling.[9] Clinicians, as advocates for patients, want to give the best possible treatment in each situation, while economists appear preoccupied with the prudent use of resources. However, clinicians are *de facto* the decision-makers deciding the allocation of resources: this is particularly the case in publicly provided health care systems where money does not follow the patient directly. With few exceptions, clinicians may defend their decisions on the

* Correspondence to: Dr James Mason, Medicines Evaluation Group, Centre for Health Economics, University of York, Heslington, York YO1 5DD, UK.
E-mail: jmm7@york.ac.uk
Source of funding: none declared.

CCC 1091–4358/99/010013–07$17.50
Copyright © 1999 John Wiley & Sons, Ltd.

basis that they believe they are providing appropriate patient care in each situation and with the information available.

Five evidence-based primary-care guidelines have recently been produced that explicitly incorporate an approach to economic thinking. These include ACE-inhibitors in the management of adults with symptomatic heart failure; aspirin for the secondary prophylaxis of vascular disease; non-steroidal anti-inflammatory drugs (NSAIDs) versus basic analgesia in the treatment of osteoarthritis; the management of dementia and the choice of antidepressants for depression.[10–14] The methodological developments, particularly with respect to incorporating cost-effectiveness concepts, are reported here and reference made to the guidelines addressing mental health topics.

A Conceptual Model

The Decision-Making Approach

A common approach to economic evaluation, the 'decision-making' approach, recognizes that those taking decisions have a range of objectives besides efficiency.[15,16] Inputs to decisions may include the decision-makers' personal values and specific notions of equity. Current thought is to provide an index of output efficiency (the cost/QALY) to contribute to the decision making process in the hope that decision makers will give such data a good weight. This assumes that the effects of various attributes of treatment decisions over time are adequately reflected in an overall single measure. The array of strong assumptions involved appear to be accepted readily by certain health economists but seldom by practising clinicians, as the literature considering the impact of cost-effectiveness studies shows little impact.[17,18,9] This may have been no bad thing since the quality of the studies themselves has often been inadequate.[19–23]

The decision-making approach assumes the existence of an audience of social decision-makers, interested in weighing the costs and benefits of treatment policy changes to all affected parties, and who will apply the results of cost-effectiveness studies. This assumption may be largely invalid in public health care systems preoccupied with politics, cost containment and process efficiency rather than health outcome efficiency. Service agreements reached at the level of medical specialities leave little scope to use an economic evaluation of an individual technology.

The rationale underpinning economic evaluation has been the belief that complex cost and benefit profiles associated with treatments can be aggregated thus handing 'an answer' to aid decision-making (at least with respect to efficiency). This has proved unproductive, in part because the methods and data have not been adequate to provide a simple answer but also because practising clinicians (the key audience) do not appear to approach individual treatment decisions in terms of economic outcomes, such as cost-utility ratios. A presumption to aggregate to a summary cost-effectiveness or cost-utility has caused many economists to lose sight of the details that affect many doctor–patient treatment decisions. All of these 'details' are potentially aspects of value that formally should be valued according to the underlying economic theory from which the QALY derives. Our hypothesis is that returning to an accurate and valid description of these effects may make the guidelines both useful to clinicians and good economics.

The Guideline Approach

The guideline development process recognizes the reality that practising clinicians are a key audience, acting as arbiters of appropriate treatment and using resources. In making decisions, clinicians balance their own preferences, those of patients and carers, patient specific information, the benefits, side-effects and safety of treatment and to varying extents (depending on the mode of reimbursement) cost. Consequently, the primary goal of guideline development is not (necessarily) to derive a cost per quality-adjusted life-year, rather the approach seeks to help the clinician to explore the profile of attributes of treatments, and aggregate these to develop well informed social preferences. The process still requires the assessment of costs and benefits of treatment to be methodologically sound, but stops at the point where the guideline members have enough information to formulate recommendations. The novel aspect is the dynamic and interactive use of economic data alongside traditional clinical inputs, in the development of clinician valuation of treatments and consequent recommendations.

Extending Guideline Methodology

Guideline Development

The composition and conduct of guideline groups, the role of the various members and the skills required has previously been reported.[2] Initially, groups are asked to define the clinical content areas of the guideline and scope of questions to be answered, ensuring a shared view between health professionals and the research team about group aims. Each guideline involves a systematic appraisal of a medical intervention in terms of the areas shown in **Table 1**. The composition of the group ensures that discussion centres on the clinical relevance of evidence summaries and practical interpretation. This then being the most current, pertinent and complete data available, each guideline sets out, or profiles, these attributes of treatment, attempts a robust presentation of uncertainties and, where appropriate, shows

Table 1. The profile of attributes systematically appraised in guidelines

- effectiveness
- tolerability
- quality of life
- safety
- health service delivery issues
- resource use
- costs in the relevant health care setting

Copyright © 1999 John Wiley & Sons, Ltd.

Table 2. Categories of evidence

Ia:	evidence from meta-analysis of randomized controlled trials
Ib:	evidence from at least one randomized controlled trial
IIa:	evidence from at least one controlled study without randomization
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

the possible bounds of cost-effectiveness that might result. The range of values used to generate low and high cost-effectiveness estimates reflects available evidence and the concerns of the guideline development group. However, simple and transparent presentation permits reworking with different values from the ones used by the guideline group. In this context, published economic analyses that adopted a variety of differing perspectives, analytic techniques and selections of baseline data are not systematically discussed in the group process. However, where differences in guideline group findings and influential or representative published economic analyses occur these are explored if helpful to the group.

Levels of Evidence and Strength of Recommendation

To assess critically information on clinical effectiveness for evidence-based guidelines, reviewers follow a process of establishing the level of evidence that individual studies provide. Papers are categorized according to study design reflecting susceptibility to bias, and questions are answered using the best evidence available. A discussion of our use of meta-analytic techniques, to summarize the results of trials, is reported elsewhere.[24] Evidence categories, shown in **Table 2**, are adapted from the US Agency for Health Care Policy and Research Classification.[25]

Recommendations are graded A to D as shown in **Table 3**. However, categories of evidence do not always simply map onto a certain strength of recommendation. First, 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 therefore attracts a lower strength of recommendation. Second, a statement of evidence may only cover one part of an area in which a recommendation has to be made, or evidence of similar quality may be contradictory. To produce comprehensive recommendations, a guideline group has to extrapolate from the available evidence. This sometimes leads to lower strength recommendations based upon category I evidence. In addition to the strength of clinical evidence then, recommendations may reflect the applicability of the evidence to the population of interest; economic considerations; guideline developers' awareness of practical issues and (inevitably) guideline developers' societal values.

Table 3. Strength of recommendation

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

To apply a strength of recommendation to cost information presents some immediate difficulties. For example, it is possible that a large well conducted trial may estimate some overall resource savings for a new treatment, but unless these findings are generalizable to normal care they may not reflect the best evidence possible. Resource measurement taken from other sources, for example insurance claims databases, may be subject to unknowable influences, particularly selection biases, and thus not provide a reliable view either. Commonly in health care, alternative treatment strategies feature small differences in outcome and precise and internally valid trials are required not just to measure differences in health outcome reliably but also (correlated) differences in resource consequences. The approach adopted by the guidelines group for cost recommendations is to apply the same categories of evidence used for effectiveness to resource use and, additionally, to establish the generalizability and relevance of findings by mapping their consequences onto current national patterns of resource use. For example the SOLVD trial, for the treatment of heart failure with an ACE-inhibitor,[26] reported rates of hospitalization for heart failure in the placebo group consistent with rates reported nationally for England. Hence the reduction in hospitalization in the active treatment group is consistent with improved health outcomes reported in the trial and plausible in the English setting. Conversely, health insurance claims databases may well show apparent overall cost equivalence (or even savings) for patients using newer and more expensive SSRI antidepressants rather than older tricyclic drugs. Selection biases (both known and unknown) may make such analyses of limited value. More fundamentally they simply may not apply to the British health care system with different approaches to, and costs of, hospitalization.

Consequently, the facets of evidence for both resource use and health outcome, as well as the generalizability of those data determine the strength of a recommendation concerning cost-effectiveness. Recommendation wording and the grade attached occurs by structured consensus, being determined by the overall quality of evidence as interpreted by the guideline development group. This approach is inevitably cautious, but may have advantages over the vagaries of clinicians enthusiasm to use new treatments, or overly enthusiastic and optimistic economic modelling.

Copyright © 1999 John Wiley & Sons, Ltd.

Applying Costs to Resource Use

While a social perspective in economic evaluation is desirable, in practice due to the (un)availability of data, analyses of costs are often limited to those borne by the NHS. Unit cost data used in guidelines are those in the public domain; it is beyond the scope of the guideline development process to conduct new costing studies.

The approach is incrementalist, thinking of the net costs and consequences of changes in practice. Costs are 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.[27] While the problems associated with average unit costs are recognized, there is no generally valid or accepted method for presenting marginal costs on items or procedures and these will vary from locality to locality. The simple presentation of analyses permits decision-makers to apply different unit costs where such information is locally available.

An Example: First Line Use of Antidepressants in Primary Care

A commentary on the recent North of England guideline addressing the first line use of antidepressants for depression in primary care is provided here. Both a shortened published form and a full resource document are available for interested readers.[14,28] The commentary is selective, illustrating the treatment of economic issues by the group.

Antidepressants are the mainstay treatment of depression in UK primary care, with one million person-years of treatment provided annually in 1995. The purchase cost of these drugs was £145 million per year but is increasing dramatically as newer (and more expensive) antidepressants receive greater use. There remains considerable uncertainty about the appropriate use of selective serotonin reuptake inhibitors (SSRIs), a relatively new group of antidepressants. SSRIs cost, on average, 5 to 6 times more to purchase than traditional tricyclic antidepressants, although the effect on overall health care costs has been hotly contested.

A Summary of the Evidence

There is good trial evidence on the short-term relative efficacy and tolerability of the various groups of antidepressants from short term trials (generally conducted over 12 weeks and in secondary care settings). Tricyclic antidepressants appear slightly more efficacious than SSRIs or related drugs, although this effect is of uncertain practical importance (Figure 1). Since the group were concerned about consistency of measurement between studies (either different instruments were used to estimate the same common underlying effect, or where poor inter rater reliability was likely) standardised scores based on standard deviation values were calculated for each trial. This has the advantage of providing a robust comparison but is hard to interpret. Calculated as a weighted

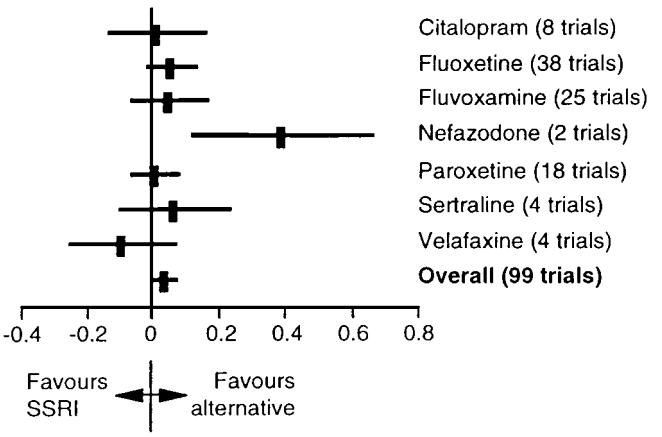

Figure 1. Efficacy of individual SSRIs and related antidepressants versus alternative antidepressant drugs: standardised weighted mean difference

mean difference, patients on tricyclics scored on average one third of a Hamilton Depression Rating Scale point more, with confidence intervals ranging from no effect to almost three quarters of a point.

SSRIs and related drugs are slightly better tolerated than tricyclic antidepressants, as measured by reduced risk of drop-out in trials (Figure 2). For one hundred patients, typical of conducted trials and over 12 weeks of treatment, four more would withdraw from treatment on a tricyclic than an SSRI, with confidence intervals from 1.5 to 6.8 Analysis of dropout by stated reason for withdrawal was not attempted because of inconsistent (and sometimes questionable) assessment in trials.

Much of the debate about the relative cost-effectiveness of SSRIs hinges on greater use of health care resources for those receiving tricyclics due to withdrawal from therapy, sedation-related accidents and toxicity associated events. The group wanted to explore the likely bounds of these effects to weigh-up the case for the use of the SSRIs. This required the use of epidemiological data since, for example, fatal poisonings associated with any antidepressant are so rare that no trial is ever likely to be conducted that will be big enough to obtain robust estimates.

Poisoning fatality data associated with antidepressants for

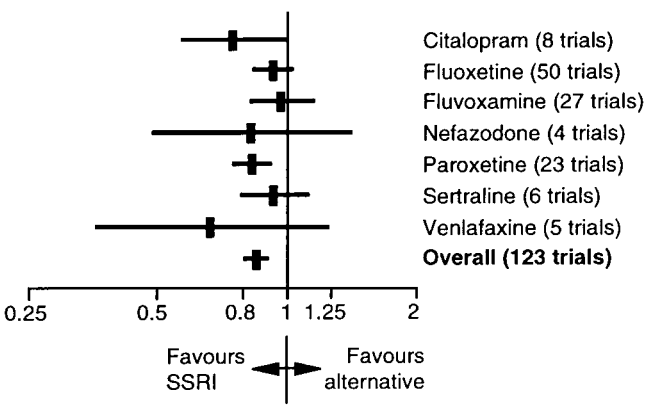

Figure 2. Relative risk of drop-out, individual SSRI or related drug versus alternative antidepressant.

Copyright © 1999 John Wiley & Sons, Ltd.

England and Wales for 3 years (1993–95) were combined with volume of antidepressant use data for the same period to estimate death rates associated with specific antidepressants. Interpreting estimated death rates requires care: higher death rates may be explained by trends in the use of certain drugs with more severely depressed and co-morbid patient groups as well as underlying pharmacological toxicity. This analysis demonstrated both the rarity of fatal overdose and the large range of rates associated with individual antidepressants. Overall, one fatality may be expected for about every 3 000 patient-years of treatment (**Table 4**). However, tricyclic antidepressants (excluding lofepramine) feature a substantial cardiovascular toxicity and a higher associated fatality rate. SSRIs, as a group, are relatively safe, with one fatality for every 100 000 patient-years of treatment. One second generation tricyclic antidepressant appears atypical: lofepramine features a fatality rate similar to the SSRIs of one fatality for every 59 000 patient-years of treatment.

Fatal poisonings are categorized according to whether single or multiple substances were ingested. Multiple ingestion involves taking other medicinal substances as well an antidepressant and so is more difficult to interpret. Nearly 70% of all antidepressant associated poisoning fatalities involved a single ingested antidepressant. Fatalities are also trichotomized as accidental, deliberate or poisoning of unknown intent. Since only 15% are recorded as accidental, it is uncertain whether fatalities could be significantly reduced by a policy of wide scale switching to less toxic antidepressants.

The majority of cases of antidepressant poisoning may not result in a fatality but be severe enough to require hospitalization. Additionally some hospitalizations due to accidents may arise from inappropriate tricyclic sedation. (Of course, sometimes sedation is an intended and necessary treatment effect.) National data exists on the total number of hospitalizations for poisoning and trauma, and it was possible to make high and low estimates of the number of these attributable to tricyclic antidepressant use.

The guideline group developed two scenarios, optimistic and conservative, to explore the merits of using SSRI and lofepramine antidepressants instead of other tricyclics. These provide a profile of the likely consequences in changes in GP and outpatient visits, and psychiatric, poisoning and accident admissions alongside the likely costs. Consequently, the incremental cost-effectiveness of using lofepramine ranged from £60 000 to £520 000 per life saved, and subsequently switching from lofepramine to an SSRI was estimated from £7.1 million to £60 million per life saved. The group consequently derived their treatment recommendations (**Table 5**).

The process of the guideline group in exploring the use of antidepressants involved much lively debate. However, when possible bounds had been explored to the satisfaction of all members, reflecting current health service resource use due to depression, the group were confident to reject claims that first line SSRI use could provide acceptable value for money. Analysis involved a top-down approach starting with all the resources in the health care system that could be assigned to depression and use of tricyclics, and examined the extent to which a policy of moving to newer drugs could reduce their use. There were inadequate data to make a sensible attempt at a bottom-up analysis.

The weakness of the modelling approach of cost-effectiveness analysis can be seen in one published analysis, which claimed to show that paroxetine was more cost-effective than imipramine[29] and which caused considerable debate.[30,31] The work of other analysts who revisited the key assumptions of the model led to opposite conclusions.[32] One large pragmatic trial conducted in the US has attempted to resolve the issue of overall health care costs using different antidepressants,[33] but, due to design limitations, its findings have no obvious interpretation in the UK setting.[34,35]

Potentially, life-years gained (a common metric in economic evaluations) could be gauged from estimates of lives saved from fatal posioning. The average age of death due to antidepressant fatal poisoning is 39 for men and 46 for

Table 4. Fatality toxicity associated with antidepressants

	Single substance death rate[1]			Single and multiple substance death rate		
	DR	ci−[2]	ci+	DR	ci−	ci+
Overall	0.339	0.328	0.349	0.492	0.479	0.505
SSRIs[3]	0.010	0.006	0.013	0.041	0.034	0.048
Lofepramine	0.017	0.010	0.025	0.062	0.047	0.076
Tricyclic and related[4]	0.577	0.559	0.595	0.808	0.787	0.830

[1]Death rate (DR) by fatal poisoning and associated with named antidepressants, per 1000 person-years of treatment.
[2]95% confidence intervals associated with each drug were calculated as DR±1.96SE, where the standard error (SE) was estimated as $(p(1\text{-}p)/n)^{\frac{1}{2}}$ using the number of treatment episodes (n) and probability (p) of a fatal poisoning. The average duration of a treatment episode was assumed to be 3 months, thus treatment episodes equal patient-years of treatment multiplied by four.
[3]Includes citalopram, fluoxetine, fluvoxamine, paroxetine and sertraline.
[4]Includes all tricyclic and related antidepressants except lofepramine.

Table 5. Selected treatment recommendations

Recommendations

- 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).
- 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).

Copyright © 1999 John Wiley & Sons, Ltd.

women, with a population average remaining life-expectancy of about 35 years for both.[36] However, this is likely to be an overestimate of life expectancy for this patient group due to comorbidity and the remaining risk of future toxic overdose. To accurately measure life-years gained would require a lifetime disease and intervention model, for which there is no adequate data. The group felt that to attempt such an analysis would be heroic, feature huge uncertainty and be unlikely to alter the final recommendations.

Discussion

Evidence-based guidelines are a vehicle for a representative group of clinicians to explore the evidence relating to a treatment area and derive appropriate recommendations that are transparent to their colleagues. In the guideline on antidepressant use there was insufficient information to sensibly calculate a cost/QALY. Even, if such calculation had been possible it is unclear how much additional value would have been attached to this information. Instead, the clinicians approached the issue of treating depression by thinking about the profile of treatment attributes. This may reflect an appropriate response to the disparate effects of treatment, some good–some bad, requiring a different cognitive model to a simple (concealing) aggregation. Seldom can a treatment's value be adequately captured by a simple cost-effectiveness construct and it is apparent that the general practitioners in the group were not working with a pre-defined notion of 'worthwhile' in the way that health-economists often approach concepts of efficiency. The cost-effectiveness approach was useful to the guideline group addressing depression to summarize a range of sequelae of treatment that they were concerned might be important once they were convinced there was no important difference in efficacy or tolerability.

Introducing economic data into evidence-based guidelines introduces some methodological challenges: specifically in providing valid, generalizable cost estimates, in the grading of cost 'evidence' and in finding a presentation helpful to clinicians. One solution, reported here, is for levels of evidence attached to the resource consequences of treatments to mirror those used for clinical effectiveness. The validity, susceptibility to bias, and generalizability of different sources of resource data are ultimately an empirical question and different viewpoints are possible. Adoption of a different system would suggest movement towards separate grading of costs and treatment effects.

The recent UK health care reforms place a great emphasis on clinical and cost-effectiveness guidelines. However, the influence of these may depend largely on whether NICE, responsible for their dissemination, is perceived by clinicians as a government organ for cost-containment or an independent body delivering valid and relevant science. A more subtle issue is whether such science can be delivered through such a centralized mechanism.

The evaluation of pharmaceuticals will remain an important aspect of guideline development. The pharmaceutical industry might consider how to enhance the likelihood of a favourable response to new pharmaceuticals from the guideline process. Most important is to ensure that forthcoming trials address the impact of their products on patient health in physical terms. Physical measures describing effectiveness, adverse events, safety, compliance and quality of life, alongside resource consequences, are most likely to inform doctor–patient interactions. Our other guideline addressing a mental health area, the primary care management of dementia, has a clear illustration of this point.[13] For the treatment of Alzheimer's disease with recently introduced donepezil hydrochloride, the group concluded on the basis of the available phase II and III trials that there was insufficient evidence to recommend its use, or to continue secondary care initiated prescribing. This finding can be contrasted with a recent Development and Evaluation Committee (DEC) report which suggested, using a cost per QALY framework, that treatment might be cost-effective.[37] The DEC finding was remarkable given the small and clinically questionable effect upon cognitive function, no improvement in quality of life in direct measurement and no evidence that the drug achieves any worthwhile effect at the stage where dementia has become severe enough to merit residential care.

A conceptual model for introducing economic concepts appropriately into guidelines has been developed. It remains a research issue whether guidelines, with or without cost-effectiveness, result in worthwhile behavioural change in clinicians, and there is an ongoing trial programme in Britain to address this. Recent published work relates only to primary care and largely addresses focused clinical questions rather than broad disease areas: the methodology needs to be applied by other teams and their experiences drawn upon. Evidence-based cost-effectiveness guidelines should be considered as a fledging science with tremendous potential, if carefully nurtured, to improve the scientific basis of clinical care.

References

1. Eccles MP, Clapp Z, Grimshaw JM, Adams PC, Higgins B, Purves I, Russell IT. North of England Evidence Based Guideline Development Project: methods of guideline development. *Br. Med. J.* 1996; **312**: 760–761.
2. Eccles M, Freemantle N, Mason JM. Methods of developing guidelines for efficient drug use in primary care: North of England Evidence Based Guidelines Development Project. *Br. Med. J.* 1998; **316**: 1232–1235.
3. NHS Executive (NHSE). Improving the effectiveness of clinical services. *Executive Lett.* 1996; EL(96)10.
4. Eddy DM. A manual for assessing health practices and designing practice policies: the explicit approach. Philadelphia, PA: American College of Physicians, 1992.
5. Institute of Medicine (eds Field MJ, Lohr KN). *Guidelines for clinical practice. From development to use*. Washington: National Academy Press, 1992.
6. Williams A. How should information of cost-effectiveness influence practice? In: Delamothe T, ed. Outcomes in clinical practice. London: BMJ, 1995.
7. Secretary of State for Health. The new NHS: modern, dependable. Cm 3807, 1997. London TSO (The Stationary Office).
8. NHS Executive (NHSE). A first class service: quality in the new NHS. HMSO, 1998. London.
9. Drummond M, Cooke J, Walley T. Economic evaluation under managed competition: evidence from the UK. *Social Sci. Med.* 1997; **45**: 583–595.

Copyright © 1999 John Wiley & Sons, Ltd.

10. Eccles M. Freemantle N, Mason JM for the North of England Evidence Based Guideline Development Project. Evidence based clinical practice guideline: aspirin for the secondary prophylaxis of vascular disease in primary care. *Br. Med. J.* 1998; **316**: 1303–1309.
11. Eccles M, Freemantle N, Mason JM for the 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. *Br. Med. J.* 1998; **316**: 1369–1375.
12. Eccles M, Freemantle N, Mason JM for the North of England Evidence Based Guideline Development Project. Evidence Based Guideline for the use of non-steroidal anti-inflammatory drugs (NSAIDs) versus basic analgesia in the treatment of pain believed to be due to degenerative arthritis. *Br. Med. J.* 1998; **317**: 526–530.
13. Eccles M, Clarke J, Livingstone M, Freemantle N, Mason JM for the North of England Evidence Based Guideline Development Project. Evidence based clinical practice guideline: the primary care management of dementia. *Br. Med. J.* 1998; **317**: 802–808.
14. Eccles M, Freemantle N, Mason JM for the North of England Guidelines Development Group. The choice of antidepressants for depression in primary care. *Family Practice* in press.
15. Williams A. Cost–benefit analysis: bastard science or insidious poison of the body politick? *J. Public Econ.* 1972; **1**: 199–216.
16. Sugden R, Williams A. The principles of practical cost–benefit analysis. Oxford: Oxford University Press, 1978.
17. Davies LM, Coyle D, Drummond MF, EC Network. Current status of economic appraisal of health technology in the European Community: report of the Network. *Social Sci. Med.* 1994; **38**(12): 1601–1608.
18. Ross J. The use of economic evaluation in health care: Australian decision makers' perceptions. *Health Policy*, 1995; **31**: 103–110.
19. Lee JT, Sanchez LA. Interpretation of 'cost-effective' and soundness of economic evaluations in the pharmacy literature. *Am. J. Hosp. Pharm.* 1991; **48**: 2622–2627.
20. Adams ME, McCall NT, Gray DT, Orza MJ, Chalmers TC. Economic analysis in randomized controlled trials. *Med. Care* 1992; **30**: 231–243.
21. Udvarhelyi IS, Colditz GA, Rai A, Epstein AM. Cost-effectiveness and cost–benefit analysis in the medical literature: are methods being used correctly? *Ann. Internal Med.* 1992; **116**: 238–244.
22. Gerard K. Setting priorities in the new NHS: can purchasers use cost-utility information? *Health Policy* 1993; **25**: 109–125.
23. Mason JM, Drummond MF. The DH register of cost-effectiveness studies: content and quality. *Health Trends* 1995; **27**: 50–56.
24. Freemantle N, Mason JM, Eccles M. Deriving treatment recommendations from evidence within randomised trials: the role and limitation of meta analysis. *Int. J. Health Technol. Assessment Health Care* in press.
25. US Department of Health and Human Services, Public Health Service, Agency for Health Care Policy and Research (AHCPR). *Acute pain management: operative or medical procedures and trauma*. Rockville, MD: Agency for Health Care Policy and Research Publications, 1992.
26. 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.
27. Mooney GH, Drummond MF. Essentials of health economics: part 1 (continued)—what is economics? *Br. Med. J.* 1982; **285**: 1024–1025.
28. North of England Evidence Based Guideline Development Project. The choice of antidepressants for depression in primary care. Newcastle: University of Newcastle, Centre for Health Services Research, 1998.
29. Jönsson B, Bebbington P. What price depression? The cost of depression and the cost effectiveness of pharmacological treatment. *Br. J. Psych.* 1994; **164**: 665–673.
30. Freemantle N, House A, Mason J, Song F, Sheldon T. Economics of treatment of depression (letter). *Br. J. Psychiatry*, 1995; **166**: 397–405.
31. Jönsson B, Bebbington P. Economics of treatment of depression [authors' reply]. *Br. J. Psychiatry*, 1995; **166**: 398–399
32. Woods SW, Rizzo JA. Cost-effectiveness of antidepressant treatment reassessed. *Br. J. Psychiatry* 1997; **170**: 257–263.
33. Freemantle N, Maynard A. Something rotten in the state of clinical and economic analyses. *Health Econ.* 1997; **3**: 63–67.
34. Freemantle N, Drummond M. Should clinical trials with concurrent economic analysis be blinded? JAMA 1997; 277: 63–64.
35. Mason JM. The generalisability of pharmacoeconomic studies. Pharmacoeconomics 1997; **6**: 503–514.
36. *Health and personal social services statistics for England* London: Department of Health, Government Statistical Service, 1996.
37. Development and Evaluation Committee (DEC). Donepezil in the treatment of mild to moderate senile demenia of the Alzheimer type (SDAT), Report No. 69. Bristol: NHS R&D, 1997.
38. Simon GE, Vonkorff M, Heilingenstein JH, Revicki DA, Grothaus L, Katon W, Wagner EH. Initial antidepressant choice in primary care: effectiveness and cost of fluoxetine vs tricyclic antidepressants. *J. Am. Med. Assoc.* 1996; **275**: 1897–1902.

Copyright © 1999 John Wiley & Sons, Ltd.