

## **Prioritisation of health technology assessment. The PATHS model: methods and case studies**

J Townsend  
M Buxton  
G Harper



---

**Health Technology Assessment  
NHS R&D HTA Programme**





**INAHTA**

### **How to obtain copies of this and other HTA Programme reports.**

An electronic version of this publication, in Adobe Acrobat format, is available for downloading free of charge for personal use from the HTA website (<http://www.hfa.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](mailto: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.hfa.ac.uk/htacd.htm](http://www.hfa.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.

# **Prioritisation of health technology assessment. The PATHS model: methods and case studies**

J Townsend<sup>1</sup>\*

M Buxton<sup>2</sup>

G Harper<sup>3</sup>

<sup>1</sup> Public and Environmental Health Research Unit, London School of Hygiene and Tropical Medicine, UK

<sup>2</sup> Health Economics Research Group, Brunel University, UK

<sup>3</sup> Centre for Research in Primary and Community Care, University of Hertfordshire, UK

\* Corresponding author

**Declared competing interests of authors:** none

Published August 2003

---

This report should be referenced as follows:

Townsend J, Buxton M, Harper G. Prioritisation of health technology assessment. The PATHS model: methods and case studies. *Health Technol Assess* 2003;7(20).

*Health Technology Assessment* is indexed in *Index Medicus/MEDLINE* and *Excerpta Medica/EMBASE*.

# 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.

The research reported in this monograph was identified as a priority by the HTA Programme's Methodology Panel and was funded as project number 94/25/31.

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.

Methodology Programme Director: Professor Richard Lilford

HTA Programme Director: Professor Kent Woods

Series Editors: Professor Andrew Stevens, Dr Ken Stein, Professor John Gabbay, Dr Ruairidh Milne and Dr Rob Riemsma

Managing Editors: Sally Bailey and Sarah Llewellyn Lloyd

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

© Queen's Printer and Controller of HMSO 2003

This monograph may be freely reproduced for the purposes of private research and study and may be included in professional journals provided that suitable acknowledgement is made and the reproduction is not associated with any form of advertising.

Applications for commercial reproduction should be addressed to HMSO, The Copyright Unit, St Clements House, 2-16 Colegate, Norwich, NR3 1BQ.

Published by Gray Publishing, Tunbridge Wells, Kent, on behalf of NCCHTA.

Printed on acid-free paper in the UK by St Edmundsbury Press Ltd, Bury St Edmunds, Suffolk.

M

## Abstract

### Prioritisation of health technology assessment. The PATHS model: methods and case studies

J Townsend,<sup>1\*</sup> M Buxton<sup>2</sup> and G Harper<sup>3</sup>

<sup>1</sup> Public and Environmental Health Research Unit, London School of Hygiene and Tropical Medicine, UK

<sup>2</sup> Health Economics Research Group, Brunel University, UK

<sup>3</sup> Centre for Research in Primary and Community Care, University of Hertfordshire, UK

\* Corresponding author

**Objectives:** To develop a method of economic evaluation and triage for research prioritisation, before the funding decision.

**Data sources:** Existing models were researched focusing on MEDLINE, HealthSTAR, IBSS and HEED.

**Review methods:** Papers of primary relevance that included a proposed model were reviewed in detail, and their models appraised using criteria adapted from the EUR-ASSESS project and the authors' previous experience. From this the PATHS model was developed. It assumes three or more possible alternative outcomes or scenarios in terms of research results: 'favourable' to the technology being assessed, 'unfavourable' or 'inconclusive'. An associated flow of benefits or disbenefits, costs or savings is identified for each potential research outcome depending on the likely implementation of the results as judged by experts. These benefits and costs are weighted and discounted in the model to give an expected incremental cost-effectiveness ratio (EICER). EICERS could be estimated for any number of research areas or proposals to inform funding prioritisation. The model was tested and evaluated on three case studies identified in liaison with the NHS R&D HTA programme and the UK Medical Research Council. These case studies were funded research projects, where full evaluation was underway and where results would be reported during the PATHS project. The studies were selected to include surgery or other invasive procedures, and non-invasive health services projects (a fourth case study did not complete during the course of the study). The three case studies included randomised controlled trials of early surgery

or observation for small abdominal aortic aneurysms, infusion protocols for adult pre-hospital care, and postnatal midwifery support.

**Results:** Each of the three assessments indicated net clinical benefit or no clinical loss of benefit, in addition to health service cost savings in excess of the cost of the trial. For two case studies, the value of the proposed trial, as evaluated by the model in the prediction, was consistent with the *ex post* evaluation, thus providing positive tests of the value of the model. In the third case meaningful *ex post* analysis was not possible as very poor compliance with the trial protocol (indicated in the *ex ante* evaluation) seriously undermined its conclusions. During the study, at the request of the UK HTA programme, the model was also applied to a funding request for a large randomised trial of  $\beta$ -interferon for multiple sclerosis treatment.

**Conclusion:** The PATHS model has a useful part to play in the research prioritisation process. Its strengths lie in its emphasis on the impact of research results on policy and practice (the keystone for NHS research) and net effects on health benefits and costs. It assesses the cost-effectiveness of the research and may identify ways to enhance the research design, endpoints relevant to implementation, analytical methods and dissemination. Further research is recommended to investigate the scope for synthesising the strengths of the PATHS model with other approaches including value of information; to compare *ex ante* and immediate *ex post* assessments of implementation with long term follow-up of actual implementation; and to assess the robustness of such approaches to the choice and number of experts used.



# Contents

<b>Glossary and list of abbreviations .....</b>	vii	Timing of the trial .....	40
<b>Executive summary .....</b>	ix	Data from experts .....	40
<b>I Background: health technology assessment prioritisation .....</b>	1	Model parameters, results and discussion .....	41
The PATHS study .....	1	Conclusions .....	42
<b>2 Review of the literature .....</b>	3	Actual results of the trial .....	43
Introduction .....	3	Testing of the model .....	43
Methods .....	3	Modelling <i>ex post</i> predicted outcomes .....	44
Commentary .....	5		
Conclusions .....	12		
Annex to Chapter 2: an application of the Townsend–Buxton model to a proposal for the evaluation of long-term use of hormone replacement therapy .....	12		
<b>3 The synthesised model .....</b>	17		
Development .....	17		
Conclusions .....	19		
<b>4 Case study 1: the costs and benefits of postnatal midwifery support .....</b>	21		
Project summary .....	21		
Timings of the trial .....	21		
Data from experts .....	21		
Model parameters, results and discussion .....	23		
Conclusions .....	25		
Actual results of the trial .....	25		
Testing of the model .....	26		
Modelling <i>ex post</i> predicted outcomes .....	27		
<b>5 Case study 2: an RCT of infusion protocols in adult pre-hospital care .....</b>	29		
Introduction .....	29		
Project summary .....	29		
Timings of the trial .....	29		
Data from experts .....	29		
Model parameters, results and discussion .....	30		
Conclusions .....	34		
Actual results of the trial .....	34		
Testing the model .....	35		
<b>6 Case study 3: MRC small aneurysm trial of early surgery or observation for small abdominal aortic aneurysms .....</b>	39		
Introduction .....	39		
Project summary .....	39		
<b>7 An application of the PATHS model: an assessment of the likely payback from a proposed UK trial of β-interferon for the treatment of multiple sclerosis .....</b>	45		
Background .....	45		
Evidence available at the time .....	45		
'Payback' analysis and recommendation .....	45		
Subsequent research and policy decisions .....	48		
Conclusions .....	49		
<b>8 Conclusions .....</b>	51		
Recommendations to HTA funders .....	53		
Recommendations for further research .....	54		
<b>Acknowledgements .....</b>	55		
<b>References .....</b>	57		
<b>Appendix 1</b> Details of the search strategy .....	61		
<b>Appendix 2</b> Trial of the costs and benefits of postnatal midwifery support .....	63		
<b>Appendix 3</b> Trial of the costs and benefits of postnatal midwifery support: questions to experts (2) .....	69		
<b>Appendix 4</b> An RCT of infusion protocols in adult pre-hospital care .....	71		
<b>Appendix 5</b> An RCT of infusion protocols in adult pre-hospital care: questions to experts (2) .....	75		
<b>Appendix 6</b> The UK small aneurysm trial .....	77		
<b>Appendix 7</b> The UK small aneurysm trial: questions to experts (2) .....	81		
<b>Health Technology Assessment reports published to date .....</b>	83		
<b>Health Technology Assessment Programme .....</b>	91		



# Glossary and list of abbreviations

## **Glossary**

Counterfactual: the situation (use of the technology) were the assessment not to take place.

*Ex ante*: before the assessment.

*Ex post*: after the assessment.

## **List of abbreviations**

AAA	abdominal aortic aneurysm	IOM	Institute of Medicine
A&E	accident and emergency	ITU	intensive therapy unit
CI	confidence interval	MOS	medical outcomes study
CMSW	community midwifery support worker	MRC	UK Medical Research Council
DRG	diagnosis related group	MS	multiple sclerosis
EICER	expected incremental cost-effectiveness ratio	NICE	National Institute for Clinical Excellence
ENBR	expected net benefits of the research	OR	odds ratio
EVPI	expected value of perfect information	PATHS	Preliminary Assessment of Technology for Health Services
EVSI	expected value of sample information	QALY	quality-adjusted life year
GHP	General Health Perception (of the SF-36)	RCT	randomised controlled trial
HEED	Health Economic Evaluation Database	R&D	research and development
HRT	hormone replacement therapy	RR-MS	relapsing remitting stage of multiple sclerosis
IBSS	International Bibliography of the Social Sciences	SF-36	Short Form 36
INF-β	beta-interferon	SP-MS	secondary progressive stage of multiple sclerosis
		TAPSS	Technology Assessment Priority Setting System

All abbreviations that have been used in this report are listed here unless the abbreviation is well known (e.g. NHS), or it has been used only once, or it is a non-standard abbreviation used only in figures/tables/appendices in which case the abbreviation is defined in the figure legend or at the end of the table.



# Executive summary

## Background

Organisations funding health technology assessment face problems of prioritisation, and some method of estimating potential returns to research is needed if limited funds are to be used cost-effectively. Most funding bodies, such as the UK NHS R&D HTA programme and the Medical Research Council (MRC), use criteria-based systems that do not include explicit calculation of cost-effectiveness and do not formally estimate returns to a research project.

## Objectives

The Preliminary Assessment of Technology for Health Services (PATHS) study aimed to develop a method of economic evaluation and triage at the stage of research prioritisation, before the funding decision. It is for use either at the stage of deciding on an area of research for funding, or at the specific proposal stage, or both, and assesses whether the additional information from an assessment will justify its cost in terms of the likely health gain and costs resulting from its impact on the use of that technology, and if so what priority should be given to that assessment.

## Method

Existing methods were reviewed against formal criteria and a model was developed that synthesised the best aspects of existing models. The approach used data from existing sources and judgements from experts, concerning possible clinical outcomes of the proposed assessment.

The PATHS model assumes three or more alternative outcomes or scenarios in terms of the research 'results': 'favourable', 'unfavourable' and 'inconclusive' outcomes. An associated flow of benefits or disbenefits, costs or savings is identified for each outcome depending on likely implementation of the results as judged by experts. These benefits and costs are discounted in the model to give an expected incremental cost-effectiveness ratio (EICER). EICERs could be

estimated for any number of research areas or proposals to inform funding prioritisation. By comparing the EICERs across research technology areas or proposals within one particular area, and the cost and effects of continuing with the current provision, a funding body could allocate funds to provide more efficient returns to research.

## Data for the model

The model is straightforward and transparent, and does not require major data collection. Data include estimates of benefits to patients, costs of the technology, level of its use in the absence of the proposed assessment (the counterfactual), likely developments in the technology during the period of evaluation, and expected changes in use of the technology given alternative outcomes of the assessment. Alternative values can be incorporated for net costs, benefits and probabilities for each scenario, and the expected level of the implementation can be adjusted, allowing the evaluation to reflect likely impact on practice as a result of reduction in uncertainty. Where available, empirical data are used, with gaps filled by expert opinion. The experts may include clinical, health economic and purchaser expertise to represent relevant decision-makers and to triangulate the estimates.

## Testing the model

The model was tested and evaluated on three case studies identified in liaison with the NHS R&D HTA programme and the MRC. These case studies were funded research projects, where full evaluation was underway and where results would be reported during the PATHS project. Two MRC- and two HTA-funded studies were selected to include surgery or other invasive procedures and non-invasive health services research projects; one case did not complete during the course of the study. The three case studies included randomised controlled trials of postnatal midwifery support, infusion protocols in adult pre-hospital care, and early surgery or observation for small abdominal aortic aneurysms.

For two case studies, the value of the proposed trial, as evaluated by the model in the *ex ante* prediction, was consistent with the *ex post* evaluation, thus providing positive tests of the model. Each of these assessments indicated net clinical benefit or no clinical loss of benefits, in addition to health services cost savings in excess of the trial cost. In the third case meaningful *ex post* analysis was impossible, as very poor compliance with the trial protocol seriously undermined its conclusions.

## **Live application of model**

During the course of the project the investigators were asked to apply the model to an application for funding a large randomised trial of  $\beta$ -interferon for multiple sclerosis treatment, submitted to the UK HTA programme. The results of this analysis illustrate further the use of the model.

## **Conclusions**

The NHS R&D programme sets relevance to the improvement of health and health services as the keystone for research prioritisation. To assess the effects on implementation the baseline level of use must be known, but this is rarely provided. Survey data may be considered an essential adjunct to a literature review, to provide a basis for assessing the relevance and potential importance of a health technology assessment, as information on the current use of a technology, and its expected trajectory, is essential to the assessment of payback. The implications are different for a new technology that would be adopted only if good evidence were provided, compared with a technology that, despite lack of good evidence, is already in use. A large part of the payback in the cases considered was due to an expectation that the research would lead to a *reduction* in the use of the technology were it proved to have low benefit. Negative results may produce high payback. An essential element of the evaluation is the explicit assessment of the counterfactual, and consideration of the length of time over which the research may influence policy. This will depend on

emerging information and changes to the technology or its competitors. In an area of rapid technological change, the policy relevance of a piece of research may be transient.

In conclusion, the PATHS model has a useful part to play in the research prioritisation process alongside existing criteria; its strength lies in its emphasis on impacts on policy and practice, and net effects on health benefits and costs. It assesses the cost-effectiveness of the research and may identify ways to enhance the research design, endpoints, analytical methods and dissemination.

## **Suggestions for HTA funders**

Applications of the model need to be conducted by competent and impartial evaluators and to be transparent. The model was tested here on primary research, but it could be applied to any form of research, including secondary analysis and reviews. Such an assessment is likely to cost £1000 to £4000, possibly more for a large or complex project. This is a small proportion of the typical research cost and should give good returns by excluding low-return proposals and improving the policy relevance of others. HTA funders should consider formal analysis of potential payback in the later stages of evaluation for projects costing, say, over £250,000. The scale and intensity of the exercise could be varied to reflect the cost, policy importance and contentiousness of the proposal.

## **Recommendations for further research**

Other developments in the literature have occurred in parallel with this work. Further research is needed:

- to investigate how to synthesise the strengths of the value of information and the PATHS approaches
- to compare *ex ante* and immediate *ex post* assessment of implementation with long-term follow-up of actual implementation
- to assess the robustness of such approaches to the choice and number of experts used.

# Chapter I

## Background: health technology assessment prioritisation

All organisations funding health technology assessments (HTAs) are faced with problems of prioritisation.<sup>1,2</sup> The continuing development of healthcare technologies has increased the number of technologies for potential assessment, while the increased demand for evidence to inform health-service practice and clinical decision-making has resulted in the rapid growth of health technology assessment programmes and proposed assessments. The funds available limit what can be evaluated: in 1997, approximately 1800 topics were proposed as candidates for assessment under the UK NHS R&D HTA programme; only 2% secured research funds in that funding cycle.<sup>3</sup> To inform this decision process health technology assessment organisations need to assess potential returns to their limited research funds and apply cost-effectiveness to the choice of trials as well as to the actual use of the technologies.

The EUR-ASSESS project (an international group designed to stimulate and coordinate developments in health technology assessment in Europe and to improve decision-making concerning adoption and use of health technology) produced a set of guidelines for the prioritisation of HTA projects.<sup>4</sup> These guidelines suggest that:

- it should be clear to all involved in a health technology assessment programme how priorities are identified
- there should be agreement regarding the approach, method and criteria for assessment between those responsible for priority setting
- the approach should reflect the goals of the programme and the resources available
- the aim of the priority setting process should be to ensure that the priorities reflect the likely costs and benefits of the possible assessments
- where possible, the method for priority setting should allow possible assessment to be rated in a systematic way against explicit criteria.

Health technology assessment organisations such as the UK NHS R&D HTA programme and the Medical Research Council (MRC) currently employ a criteria-based system for the prioritisation of research projects.<sup>1,2</sup> These systems

do not include explicit calculation of the cost-effectiveness of undertaking research. Those advising research funding decisions are typically given a checklist of criteria (e.g. Would the project lead to a reduction in uncertainty? Are the costs justified? Are the objectives clearly stated?) against which research projects are judged. Funding bodies weigh likely costs and benefits, but usually in general and qualitative terms. They rarely attempt to estimate systematically the cost-effectiveness of undertaking a research project. If methods were available to provide ready and reliable quantitative estimates of the likely returns from competing uses of their resources, health technology assessment funding organisations would have greater reassurance as to the appropriateness of their funding decisions.

### The PATHS study

For ease of reference, this study is called the PATHS (Preliminary Assessment of Technologies for Health Services) study. It was set up to develop a practicable method for economic evaluation and triage at the stage of prioritisation of health technology assessment either at the stage of considering general topics or at the stage of specific proposals, but before the funding decisions.<sup>5</sup>

The study aims to clarify whether it is possible *ex ante* to identify whether the costs of a health technology assessment are justified in terms of its likely implication for adoption or reduced use of that technology, and how to prioritise those projects that appear to be justified. Existing methodological approaches are reviewed and a formal methodology is synthesised. This model is tested with the aid of expert panels in three case studies of evaluations funded and underway as the research commenced and expected to report during the course of the project.

The proposed approach requires estimates of parameters from existing data and professional judgements concerning:

- examples of possible clinical outcomes of the proposed assessment

- benefits to patients from these clinical outcomes
- estimates of costs of the technology to be assessed
- the direction of policy or use of the technology in the absence of the proposed full evaluation
- likely developments in the technology during the lifetime of the assessment
- appropriate policy changes in use of the technology for possible clinical outcomes of the assessment
- actual likely policy changes.

These parameter estimates need to be made with as full information as possible. Some estimates may be available from existing data, but some will not be. In this study a small expert panel was formed for each case study, typically consisting of two clinical or scientific experts in the field, a

health economist and a ‘purchaser’ or decision-maker, all of whom would be independent of the proposal. Possible options for the parameters were discussed with each expert and these assumptions fed into the model with cost and benefit data for different outcomes, the cost of the technology assessment and an agreed period of payback.

Results from this ‘*ex ante*’ evaluation were tested against the actual outcomes from the technology assessments and the implications for policy and practice discussed with the panels. By focusing on a sample of funded technology assessments, the aim was to obtain an indication of whether such a formal evaluation process might have added useful information to the funding process and possibly have led to a different decision.

# Chapter 2

## Review of the literature

### **Introduction**

In recent years the evaluation of the returns to public sector investment in research and development has moved up the policy agenda.<sup>6–8</sup> The reasons for this include the need for increased accountability, improved cost-effectiveness of research and an improved information base for decisions about allocating research funds. Much of the initial work and most of the pioneering current work has been developing methodology and applying it retrospectively to completed research; that is, looking at payback.<sup>9,10</sup> These concerns for accountability and cost-effectiveness have extended the discussions to management<sup>6,11,12</sup> and to monitoring.<sup>13–15</sup> Work on *ex ante* assessments, attempting to predict returns to research before funding has been granted and the research undertaken, has been rare, although there is increasing interest in the role that economic analysis can play in the prioritisation of research funding.

### **Methods**

#### **Search methods**

The primary aim of this literature review was to identify relevant published studies that proposed a conceptual/theoretical model for relating prospectively the costs, at any level, of undertaking a health-services research project (or programme) with the projected benefits (however measured) of that research, which could be used to prioritise non-commercial research topics or proposals. The secondary aim was to note more general papers concerning the assessment of payback to research, which provide context and additional ideas for the discussion below.

Given the difficulties of undertaking systematic searches on methodological issues, particularly as here, where the known relevant literature uses a variety of non-specific terminology, search strategies were developed to achieve higher sensitivity rather than specificity. The strategy used a variety of possible search terms, to maximise the yield of relevant studies. The search focused on the following electronic bibliographic databases: MEDLINE, HealthSTAR and IBSS

(International Bibliography of the Social Sciences). The search was limited to English language journals only. Articles identified by the searches were reviewed for relevance by two of the team (GH and JT) on the basis of title (and abstract where available). If relevant, or in the case of doubt, full articles were obtained and reviewed. In addition, reference lists of relevant articles were scanned for further possibly relevant studies.

The initial search was carried out in May 1998, with a subsequent updating search of the literature in January 2002 (with a review conducted in an equivalent manner by MB and JT) to inform the concluding chapter of parallel developments in the literature. The updating search was extended to cover the Health Economic Evaluation Database (HEED): this did not identify any papers of primary relevance that had been missed by the original search in 1998.

Further details of the specific search terms are given in Appendix 1.

#### **Appraisal criteria**

Papers of primary relevance that included a proposed model were reviewed in detail, and their models appraised using criteria adapted from the EUR-ASSESS project and previous experience of the authors.<sup>4,5,16</sup>

The appraisal criteria used were as follows.

- Is the aim of the model clear? Is it concerned with identifying technologies for assessment or for prioritising individual trials, or both?
- Does the model consider alternative outcomes of a trial, and the uncertainty surrounding the outcomes in a probabilistic framework?
- Does the model consider the counterfactual – what would happen if the trial were not funded?
- Does the model consider the implications of the likely implementation of an intervention following the reporting of the trial's conclusions?
- Could the model be operationalised (i.e. would a data source be available)?
- Are other forms of payback to the research considered? In particular, are weights assigned to a range of potential benefits to research?

- Could the model be used within an iterative economic evaluation framework, using data from earlier stages and also informing later stages?
- Is the assessment process transparent and easily interpretable by a research funding body and potential funding applicants?
- Are the uncertainty and subjectivity related to the data collection and modelling process considered?
- Is the process likely to be efficient? In particular, would the process itself consume sizeable resources such as researchers' time and costs of data collection?
- Could the process be flexible to adjust to different scales of research?
- Could the model accommodate alternative outcome measures of a trial if necessary?
- Could the model accommodate analysis of particular subgroups of patients?

## **Results**

The searches identified 799 possible separate articles. The full text of 26 of these was obtained. From the full texts, it emerged that ten papers related to seven published models that fulfilled the primary requirements for inclusion (a proposed model of research assessment or prioritisation, or a quantitative example of assessing the benefits of a research project in relation to costs).<sup>16–25</sup> Other papers did not include an appropriate model but some had relevant material that informed the analysis and critique of the models, and are referred to subsequently in discussion.

The seven major models are described in brief below and then appraised using the predetermined criteria.

### **1. Claxton and Posnett<sup>17</sup>**

This model assesses the value of new information using a Bayesian approach. The model incorporates the expected value of perfect information (EVPI) and the expected value of sample information (EVSI), expressed in monetary units. It is a 'two-hurdle' model designed to make an assessment first at the technology level and then at the specific research proposal level. The EVPI is used first to assess whether a reduction in the uncertainty surrounding a technology area would be valuable. The criterion set by Claxton and Posnett is for the EVPI to exceed the fixed costs of the research. To ensure that the design of a specific trial is cost-effective when conducted at optimal size, it must clear the second hurdle, that the expected net benefits of the research (ENBR) are greater than zero, at optimal sample size.

### **2. Detsky<sup>18–20</sup>**

This model, first set out by Detsky in 1985, relates the impact of a trial, with given power and significance level, to the distribution of the risk reduction of the problem under consideration.<sup>18</sup> For example, this could be in terms of the number of people currently dying per year following treatment with a traditional therapy compared with the situation following the new assessed therapy. By relating the risk reduction to the trial cost, he estimates the cost-effectiveness ratio of the research, assuming 100% implementation following the conclusion of the trial and a theoretical distribution of both risk and costs. His subsequent papers apply this model (retrospectively) to seven major randomised controlled trials (RCTs) that had been recently completed and reported.<sup>19,20</sup>

### **3. Drummond and colleagues<sup>21</sup>**

This decision-analytical model was developed to assess retrospectively the returns to a specific applied research project, the Diabetic Retinopathy Study, funded by the National Eye Institute from 1972 to 1981. Adjustments can be made to the probabilities within the decision tree (the relative probabilities of whether a patient has diabetic retinopathy before and after the trial) and to the size of the population benefiting from the intervention following the trial, as well as the benefits they might receive and the costs of the implementation. Although applied retrospectively here, the same framework could be used prospectively.

### **4. Eddy<sup>22</sup>**

The Technology Assessment Priority Setting System (TAPSS) models the effect of a range of factors on costs and benefits resulting from a technology assessment. It is a decision-analytical model and factors include the number of potential candidates for the intervention under evaluation, the probability that the assessment will arrive at a particular result, the implementation following that result and the population likely to receive the technology. The model systematically allows for variation in these key variables, and the uncertainty in the estimation of the variables. The model can incorporate different possible outcomes and can combine alternative results probabilistically into a single figure. It forms the basis of a number of later decision-analytical models (e.g. Townsend and Buxton<sup>16</sup>) and was an important influence on the prioritisation process developed for the US Agency for Health Care Policy and Research by a Committee of the Institute of Medicine in 1990 and 1992.<sup>1,26</sup>

### **5. Phelps and Parente<sup>23,24</sup>**

Phelps and Parente developed an uncertainty model that estimates the implied value of assessing a technology based on the unexplained variation in medical expenditure. It presented an application comparing procedures based on data on variations in expenditure between different counties within New York State.<sup>23</sup> The model assumes that all variation in expenditure not explained by demographic or socioeconomic factors is due to the uncertainty about the outcome of the intervention. It makes a further assumption that the 'correct' level of expenditure is approximated by the adjusted mean level of expenditure, and that there is a fixed elasticity of demand for all health technologies. Welfare losses associated with these unexplained variations in expenditure are estimated and compared across diagnosis related groups (DRGs) to identify the estimated highest payoff from potential research spending. The model is focused at the technology rather than the research proposal level. It is an imaginative and operational concept, but is limited both by its strong assumptions and by its broadbrush approach of associating just one technology assessment with each DRG. Small corrections to the computations in the original paper were reported subsequently by Phelps and Mooney.<sup>24</sup>

### **6. Townsend and Buxton<sup>16</sup>**

This model, which also uses decision analysis, was developed and applied to a proposed trial of the long-term use of hormone replacement therapy (HRT). It is an *ex ante* model for use at technology or project level and explicitly considers three (or more) broad scenarios of potential trial outcomes: positive, negative and inconclusive. Changes in the long-term use of the technology are predicted for each outcome scenario, and the net effects of these changes estimated in terms of costs and benefits. Discounted costs and benefits are weighted to obtain estimates of the overall cost-effectiveness of the proposed trial. The authors were able to draw on detailed empirical data and published estimates of the short- and long-term benefits of HRT. For other technologies, such detail might not always be available, but the authors argue that expert opinion could be used to provide best estimates. An application is given in the annex at the end of this chapter.

### **7. Weinstein<sup>25</sup>**

This, the earliest of the models, focused on a specific choice between research on a bioassay for a potentially carcinogenic, but widely used, chemical and a prospective study of the cancer-

reducing effects of dietary β-carotene. It provides a generalised analytical framework and a brief summary of somewhat simplified applications of that framework to the specific contexts. Essentially, it is based on standard principles of decision analysis and cost-effectiveness analysis, in a model that recognises the inherent uncertainty and (at least in principle) a series of possible outcomes from the research, each assigned a prior probability. It explicitly recognises the need to make a judgement about the extent to which policy and/or practice will follow the results of the research and so the extent to which potential benefits will be achieved. Its conceptual framework has been broadly adopted (apparently without specific recognition) by Detsky, Eddy, Drummond and colleagues, and Townsend and Buxton, and it also provides an early reference to the EVPI developed by Claxton and Posnett.

In *Table 1* the seven models are appraised against the set criteria given above.

Three of the papers identified in the search focused on the limited issue of determining the appropriate sample size for and design of studies, in some cases linking the strength of evidence to likely impact.<sup>27-29</sup> The issues these raised are important, and were noted in our thinking about the formal models and their use.

## **Commentary**

A number of issues arose from reviewing the papers. Although there was a wide literature alluding to the issue of prioritisation, considering the process by which research funds are allocated or addressing the effect of trial design (particularly sample size) on the likely impact of the study, the search identified only seven models that aimed to determine systematically whether a research area or proposal was likely to be a cost-effective investment.

## **Aims of the models, their methods and levels of information**

The criteria used to assess the models are based on how practical they are in terms of effectiveness and efficiency for general use. In general, the models had similar aims, namely to assess the 'value' of undertaking a particular piece of health-services research or health technology assessment and, although the precise methods used vary in each case, there is a great deal of similarity (not explicitly recognised or acknowledged and probably independently developed) in the papers.

**TABLE I** Appraisal of the seven models

	<b>Claxton and Posnett (1996)<sup>17</sup></b>	<b>Detsky (1985, 1989, 1990)<sup>18-20</sup></b>	<b>Drummond et al. (1992)<sup>21</sup></b>	<b>Eddy (1989)<sup>22</sup></b>	<b>Phelps and Parente (1992)<sup>23</sup> and Phelps and Mooney (1992)<sup>24</sup></b>	<b>Townsend and Buxton (1997)<sup>16</sup></b>	<b>Weinstein (1983)<sup>25</sup></b>
Are the aims of the model clear? Is it concerned with identifying technologies for assessment or for prioritising individual trials, or both?	Yes. The model's aim is to test the efficiency of a trial design and promote consistency in decision-making and priority setting between research and service provision	Yes. The model relates the cost of a trial to its size and to clinically important benefits identified in the trial, to assess the marginal productivity of research funding, when expanding or contracting trial size	Yes. The objective is to develop and test a methodology for assessing the social costs and benefit of research. It is designed to consider prospectively whether a research project should be funded	Yes. The model developed was 'a simple framework and quantitative method for estimating the value of assessing different technologies'. It considers a proposed assessment and evaluates how it will affect health and economic outcomes	Yes. It uses the unexplained variation in medical expenditure by DRG between New York counties, to estimate an implicit value of assessing technology, to that related DRG	Yes. It develops a methodology for <i>ex ante</i> evaluation of proposed technology assessment, comparing net costs and benefits from use of the technology following an assessment's conclusion with the likely practice in the absence of the assessment	Yes. The paper sets out a quantitative approach to setting priorities between competing uses of research funds using examples relating to cancer. It uses a cost-effectiveness framework to assess cost per cancer death/cost per life year from specific proposed trials
Does the model include alternative outcomes of a trial?	Yes, to an extent. The model compares the cost of a trial with the expected net benefits for different sample sizes. Alternative outcomes are incorporated in the determination of the expected benefits	Yes. To calculate the power of the trial, the model uses an expected level of risk reduction, taking all possible risk reductions (-100% to +100%) into account by weighting them with the probability that each outcome will occur	Yes. The model is constructed using a decision-theoretic approach, with the alternative outcomes of the trial represented by the probabilities assigned to the different events with sensitivity analysis applied to these probabilities	Yes, the alternative conclusions of the assessment are reflected in the range of delta results selected for the assessment. For each delta result (the result of an assessment that can potentially change the use of a technology) a probability is estimated that reflects the likelihood that the assessment will reach that delta result	No. There is a single value for each DRG	Yes. Alternative outcomes are considered explicitly, with positive, negative and inconclusive outcomes defined in terms of varying health gains or losses and weighted with the probabilities of the trial reaching those outcomes	The general model allows for a range of research outcomes (in terms of levels of effect), although this potential is not used in the cases presented in the paper. It does use a prior probability of the effect

continued

**TABLE I** Appraisal of the seven models (cont'd)

	Claxton and Posnett (1996) <sup>17</sup>	Detsky (1985, 1989, 1990) <sup>18-20</sup>	Drummond et al. (1992) <sup>21</sup>	Eddy (1989) <sup>22</sup>	Phelps and Parente (1992) <sup>23</sup> and Phelps and Mooney (1992) <sup>24</sup>	Townsend and Buxton (1997) <sup>16</sup>	Weinstein (1983) <sup>25</sup>
Can the model be used within an iterative economic evaluation framework, using data from earlier stages and also informing later stages? i.e. at technology and project level?	Possibly. The calculation of the EVSI and the EVPI may be of limited value in the later stages of economic evaluation	Possibly for trials of different sample size	Potentially by adaptation although developed for a specific trial	Potentially, as the detailed calculations involved in the TAPSS model would be of benefit in identifying important subgroups for specific analyses during later stages of the assessment	No, as it is a specific macro model	Yes. Easily adaptable to either stage	Possibly, although conceptualised in terms of a trial (a 'test' of a hypothesis) with specific sensitivity and specificity
Is the process transparent and easily interpretable by a research funding body and funding applicants?	Not easily. It is a theoretical model, using techniques and concepts not easily interpreted by potential research applicants	Not easily. The concept of the prior distribution of the risk reduction is complex and subjective, and may not be fully understood	Yes. The decision-analytical approach clearly shows the processes of the evaluation, the assumptions made and what adjustments have been made	Uncertain, as the model involves the interaction between many factors. The calculation may not be easily understood	No. As a broad field methodology with strong assumptions about methodology and interpretation it is neither transparent nor intended for use at project level	The process is transparent and fairly easy to interpret	The published account is rather brief and lacks transparency, but it could easily be presented in greater detail, resulting in greater transparency and interpretability
Does the model consider the counterfactual – what would happen if the trial were not to be funded?	Not explicitly	No	Yes	No	The methodology centres on local practice moving directly to current average	Yes	Does not specifically allow for counterfactual change compared with current position
Does the model consider the likely implementation of the intervention following the reporting of the trial's conclusions?	Not explicitly	No	Yes	Yes	It is implied to be 100% immediately	Yes	Yes. It makes a specific judgement as to the extent to which the desired change would occur

continued

**TABLE I** Appraisal of the seven models (cont'd)

	Claxton and Posnett (1996) <sup>17</sup>	Detsky (1985, 1989, 1990) <sup>18-20</sup>	Drummond et al. (1992) <sup>21</sup>	Eddy (1989) <sup>22</sup>	Phelps and Parente (1992) <sup>23</sup> and Phelps and Mooney (1992) <sup>24</sup>	Townsend and Buxton (1997) <sup>16</sup>	Weinstein (1983) <sup>25</sup>
Is there a data source for the model to be operationalised?	Theoretical mode. No example used	Published literature	Published literature and expert opinion	Estimated by author	Published literature and estimates based on existing US county data	Published literature and author estimates	Published literature and subjective estimates by author
Are other forms of payback considered? If so, what and how?	Other forms of payback are not explicitly considered.	No	No. Health gains only	No, but model can be re-estimated using other outcomes	No	No	No
Are weights assigned to different potential benefits to research?	Benefits are based on patient utilities, converted into population benefits						
Are the uncertainty and subjectivity related to the data collection and modelling process considered?	Not explicitly, although the EVPI could be recalculated, with adjustments in the probabilistic framework	Yes, through sensitivity analysis around the prior distribution of the risk reduction and the power curve	Yes, through sensitivity analysis in calculating the cost-effectiveness ratios	The model could be recalculated using different values for the variables	No, although it could be introduced	To an extent, through adjusting the different factors involved following a trial's conclusion	Yes, to some degree through sensitivity analysis
Is the process likely to be efficient? In particular, would the process itself consume sizeable resources such as researchers' time, costs of data collection?	Uncertain. Complex calculations	Process is relatively simple, but subject to the interpretation of the prior distribution	Modelling exercise would be relatively straightforward	The data required to calculate the result require either estimation or resource-intensive collection	Interesting approach, but the categories are probably too broad to be useful. Process relatively simple and based on existing data	If data or informed opinion available, then yes. Costs could vary according to the level of detail used	Process is relatively simple and allows for use of informed opinion

continued

**TABLE I** Appraisal of the seven models (cont'd)

	<b>Claxton and Posnett (1996)<sup>17</sup></b>	<b>Detsky (1985, 1989, 1990)<sup>18-20</sup></b>	<b>Drummond et al. (1992)<sup>21</sup></b>	<b>Eddy (1989)<sup>22</sup></b>	<b>Phelps and Parente (1992)<sup>23</sup> and Phelps and Mooney (1992)<sup>24</sup></b>	<b>Townsend and Buxton (1997)<sup>16</sup></b>	<b>Weinstein (1983)<sup>25</sup></b>
Could the process be flexible to adjust to different scales of research?	Uncertain: not applied to data	Yes: the model was applied retrospectively to a number of trials ranging in cost from \$78,000 to \$150,000	Yes, although the application was to the NEI Diabetic Retinopathy Study, the cost of which was \$10.5 million. For an application to a smaller trial, the sensitivity analysis would need to be more precise	Data on small trials unlikely to be available to the extent required by the TAPSS model	No	Potentially more applicable to big trials, such as the proposed HRT trial to which the model has been applied. Implementation following trial outcomes will require precise sensitivity analysis for smaller trials	Does not obviously lend itself to small studies
Can the model accommodate alternative outcome measures of the trial, if necessary?	It can accommodate a number of patient-based outcomes, as the effectiveness is based around patients' utilities. However, it may be difficult to include other forms of benefits	Difficult, as effectiveness is based on change in risk reduction, requiring one health-related outcome	Use of QALYs is possible, which combine a number of health-related outcomes. However, other forms of benefits are not considered	Yes, but the model would need to be recalculated and the results combined with some weighting procedure	All outcomes subsumed as potential benefits are based on the assumed welfare losses from deviation from current average practice	Yes, application can use QALYs, combining a number of health outcomes	Example estimates costs per cancer death avoided, per life saved, and per year of life saved
Can the model accommodate the analysis of particular subgroups of patients?	Not explicitly considered, but would be possible	Not explicitly considered, but would be possible	Yes, within the decision-analysis framework	Yes	No, it is relevant to average county practice	Not explicitly considered, but would be possible	Not explicitly considered, but would be possible

It could be argued that conceptually all the models can be logically linked back to that of Weinstein, except that of Phelps and Parente which builds on the work relating to local and regional variation in the use of technologies pioneered by Wennberg.<sup>30</sup> This latter approach can only be used in relation to research to reduce uncertainty concerning existing technologies, and the quantification of benefit rests on the heroic assumption that on average current practice is desirable. All the other more satisfactory methods are some variant of a decision-analytical approach, seemingly first set out by Weinstein, sometimes with the costs and benefits for each trial outcome and the subsequent implementation pattern combined probabilistically. Such an approach has the advantages that the benefits of an assessment are clearly defined in terms of the effect of the health technology assessment on practice, and as a result of health outcomes.

Those models that require a formal calculation of the expected value of the information provided from an assessment, such as Claxton and Posnett,<sup>17</sup> pose two major problems. The level of information available before the trial would need to be combined into a prior distribution, for example, of the perceived mortality risk reduction of undergoing a surgical procedure; this may be difficult to carry out and is not likely to be transparent. Also, some of these models attempt to convert the value of increased information into financial terms; the problems of this have long been recognised. As Buxton and Hanney argue, such problems are “probably no more readily soluble for research payback assessment than for the assessment of the interventions themselves.”<sup>10</sup>

### **Staged approach**

A major test of the efficiency of a model is likely to be the ease with which the model could fit into a health technology assessment funding organisation’s existing processes of funding allocation. Sculpher and co-workers described the possibility of using economic evaluation iteratively as part of the health technology assessment process,<sup>31</sup> possibly with earlier preliminary analysis at technology level being used to inform later decisions, at project level or in identifying subgroups on which the full evaluation might concentrate. The earlier stages of the iterative process involve the prioritisation of technologies. Data collected for analysis at this stage might be used productively at a later stage. Buxton and Hanney<sup>9</sup> described a three-stage approach to the prioritisation in terms of estimating the payback from research. The first two stages (theoretical

potential payback and expected potential payback) consider topics or technologies, while the third stage (likely realisable payback) considers specific research proposals designed to reduce specific elements of uncertainty surrounding the technology that had been identified in the previous two stages. The more theoretical models may not have the requisite characteristics to satisfy this overall three-stage approach.

### **Dealing with uncertainty**

It is uncertainty that prompts the initial demand for some form of health technology assessment.<sup>32-34</sup> The nature of this uncertainty is most appropriately investigated at the technology stage, to see how detrimental it is and, if ameliorated, whether net benefits might accrue. The extent of the uncertainty is important, and other things being equal, the more ‘unknowns’, the higher the likely cost of the necessary specific research projects. The cost of a specific proposal will depend on what it is we are uncertain about. Uncertainty may surround the cost of a particular intervention, the effectiveness of a particular intervention, or both the costs and the effectiveness. Identification of the nature and extent of the uncertainty will influence both the research proposal and the full evaluation stages. There will also be uncertainty and some degree of subjectivity about implementation; that is, how the research results may influence practice. This will depend on the effectiveness of existing means of dissemination of information among practitioners, commissioners and policy makers, and on the professional and institutional rigidities. There will also be uncertainty surrounding the data used; the models could be recalculated with sensitivity analysis.

Models, such as those of Drummond and colleagues<sup>21</sup> and Townsend and Buxton,<sup>16</sup> that use direct health outcomes within a probabilistic, decision-analytical framework, and clear measures of the proportion of the population likely to benefit from the intervention following the trial’s conclusion, are likely to be able to handle uncertainty most systematically. The more ‘theoretical’ approach of Claxton and Posnett<sup>17</sup> does not focus attention on the uncertainty surrounding variables such as the extent of implementation. In our context, the major criterion of the efficiency of the model is how likely it will be able to influence decisions made by a research-funding organisation. This may raise problems for models using concepts that may appear abstract to potential funders or funding applicants.

## **The incremental effects of a trial: considering the counterfactual**

To consider the incremental effects of the trial on subsequent practice, a model needs to include what practice would be were the trial not funded. It may not be sufficient to assume that current practice continues. Many assessments take place against a background of changing practice. The incremental effects of the trial relating to changes in practice need to be explicitly accounted for in terms of health gains. A trial that finds in favour of an intervention that would have been introduced or continued in use anyway, but would not have prevented its adoption had it found against the intervention, is clearly unnecessary.

## **The likely impact on practitioners' behaviour**

We take the view that given that the ultimate purpose of health technology assessment is improved healthcare, the outcome needs to include the implications of the resultant likely use of the technology, and so it would be insufficient for outcomes to be expressed in terms of the reduction in uncertainty alone; that is, an assessment's health-related benefits must consider the extent to which an assessment's conclusions will affect health-service practitioners' subsequent behaviour. A number of models, including those of Detsky and Phelps and Parente, assume full implementation of the intervention following the report of a trial's findings. Organisational barriers, ineffective information dissemination and poor management of change are likely to operate at varying levels, and prevent or delay the immediate adoption of the more effective or cost-effective technology or intervention in some organisations. A realistic model would allow some means of assessing the extent of the likely adoption of the more cost-effective technology, perhaps through the use of expert opinion.

## **Non-health benefits**

A conflict arises between the multidimensional nature of payback from health-services research and the need for a precise quantitative model. Buxton and Hanney<sup>9</sup> identified five different categories of potential benefits from health-services research: knowledge benefits, benefits to future research, political and administrative benefits, health benefits and broader economic benefits. Health technology assessment organisations, according to their objectives and sources of funding, may place different emphases on these, and it is possible to argue that a model designed to prioritise research projects should relate these potential benefits from research

directly to the resource implications of undertaking the research. As stated earlier, such a model would require estimates of the relative importance of each form of benefit. For some health technology assessment programmes, it is clear that the prime motivation is to improve the cost-effectiveness of the health system. Any other form of benefit or payback, albeit important, would be considered secondary. Therefore, preliminary cost-effectiveness studies using quantitative models to evaluate a research proposal have understandably concentrated on the health benefits, to reduce subjectivity in terms of weighting alternative forms of payback from research. Other forms of benefits, however, should not be ignored and, whereas the majority of applications of the models concentrate on the health benefits, minor modifications of models such as TAPSS and the other decision-analytical models would allow other forms of payback to be assessed, and non-health benefits could and should be considered alongside the model.

## **Resource requirements to operationalise the model**

Central to all these considerations is the availability of data required to operationalise a model. A number of models rely on existing published data. Frequently, given the inherent uncertainty, appropriate data may not be available and a surrogate alternative, such as expert opinion, may be necessary. Data requirements of a model may be too demanding or too costly for a modelling exercise to be efficient. Sheldon<sup>35</sup> described modelling as a way of "reducing complexity to simple elements in order to better understand the way a system works or to predict its effects." However, if the model itself requires data collection that is likely to consume significant resources, as might for example the TAPSS model, the cost of the preliminary modelling exercise might begin to approach the costs of undertaking the assessment itself, in which case the efficiency of the model and its usefulness to health technology assessment funding organisations come into question. Many of the models lend themselves best to situations where cost-effectiveness modelling of the intervention has been previously undertaken.

## **The timing of the costs and benefits of the technology**

The timing of the costs and benefits relating to the health technology assessment will affect its overall cost effectiveness.<sup>36-38</sup> It is unlikely that the payback, and the costs of changes in implementation, will occur immediately after the trial.<sup>39</sup> All the models considered above are able to

incorporate appropriate discounting; however, a number of the applications on which the models were tested did not do so. Related to this is the potential lifespan of a technology. If a particular technology or pharmaceutical considered for assessment is expected to be superseded shortly by a rival that is at an earlier stage of development, then the shelf-life of the technology being considered for assessment might be relatively short and a substantial investment in the assessment may not be efficient.<sup>40</sup> The benefits of the research over a given period need therefore to be adjusted for varying shelf-life. The value of an assessment may also be reduced or negated if results of a similar or related assessment of the same technology are pending.

## Conclusions

Several useful models for prioritisation of technology assessment have been set out in the published literature. They share many common features and cannot be seen as conceptually independent. None, as formally presented, fully met the criteria adopted. The most readily operationalisable models, such as those of Eddy, Drummond and co-workers, and Townsend and Buxton, are based on decision analysis. The last of these offers an easily applied model making explicit use of weighted alternative options. Its approach is demonstrated in summary in the annex at the end of this chapter. The Drummond model has the advantage that it uses expert

opinion as well as subgroup analysis. A synthesised model is needed that draws on the best of these existing models.

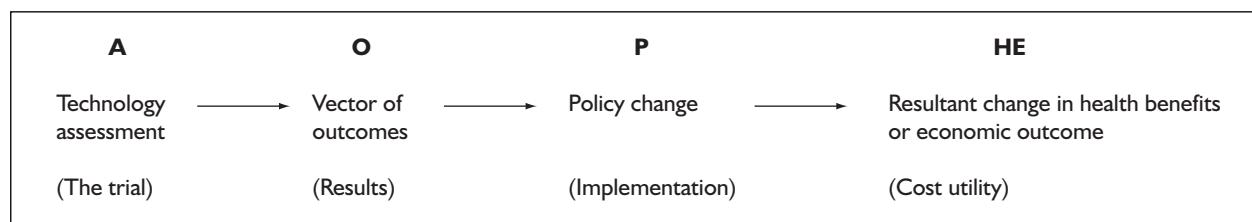
## Annex to Chapter 2: an application of the Townsend–Buxton model to a proposal for the evaluation of long-term use of hormone replacement therapy

The Townsend–Buxton model was developed in response to a request by the UK MRC and Department of Health to evaluate the likely payback to an extensive long running trial of HRT that they were considering for funding. A summary of the application of the model is given here.

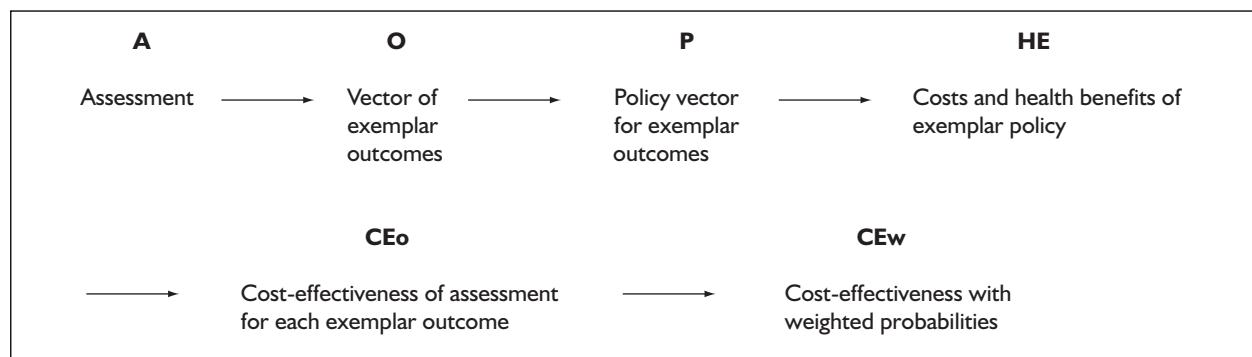
### The model

The model is based on a sequential flow of assessment, outcome, policy change, and health benefits and costs (*Figure 1*).

Each outcome along a spectrum of possible outcomes will have policy implications with implications for different benefits, risks and costs. For simplicity this spectrum of outcomes may be narrowed. These may typically form a triage of positive, negative or uncertain, and the scenario analysis would form a sequence as shown in *Figure 2*.



**FIGURE 1** Assessment/policy sequence (A, assessment; O, outcome; P, policy change; HE, health benefits and costs)



**FIGURE 2** Scenario analysis sequence (A, assessment; O, outcome; P, policy change; HE, health benefits and costs; CEo, cost-effectiveness of outcome; CEw, cost-effectiveness with weighted probabilities)

The overall value of the trial would relate to the stream of costs and benefits flowing from the projected exemplar change of policy, the costs and benefits of the counterfactual (status quo) and the costs of the assessment itself, all discounted to a baseline as given in equation (1):

$$\text{CE}_0 = \frac{\left[ \sum_t \sum_g \text{DCP}_{otg} - \sum_t \sum_g \text{DCNP}_{otg} + \text{DCA} \right]}{\left[ \sum_t \sum_g \text{BP}_{otg} - \sum_t \sum_g \text{BNP}_{otg} \right]}$$

where  $\text{CE}_0$  is the cost-effectiveness for outcome  $o$  and associated policy change  $p$ ;  $\text{DCP}_{otg}$  is the discounted cost of policy change  $p$  for outcome  $o$ , year  $t$ , group  $g$ ;  $\text{DCNP}_{otg}$  is the discounted cost for the counter factual outcome  $o$ , year  $t$ , group  $g$ ; DCA is the discounted cost of the assessment itself;  $\text{BP}_{otg}$  is the benefit from policy change  $p$ , outcome  $o$ , year  $t$ , group  $g$ ;  $\text{BNP}_{otg}$  is the benefit given no policy change, outcome  $o$ , year  $t$ , group  $g$ .

### **The proposed trial scenario analysis**

Research over many years had suggested that the major long-term effect of HRT was a likely reduction in risk of cardiovascular disease, stroke and osteoporosis, principally coronary heart disease.<sup>41</sup> An RCT was proposed by the UK MRC and was set to identify a 25% reduction in ischaemic heart disease and stroke event rate within 10 years for women treated with HRT for 10 years, compared with women on placebo for the same time, with a power of 80% at a significance level of 5%. Collaboration from research centres in Europe and Australia was proposed to raise the power of the trial.

There was evidence that HRT also reduced the risk of osteoporosis,<sup>42,43</sup> but increased the risk of breast cancer.<sup>44</sup> The sample size for the proposed trial was set on the cardiovascular events, but other secondary end-points were included relating to osteoporosis (hip fractures and vertebral fractures) and breast cancer. The risk of hip fracture occurs at an older age than cardiovascular disease, and the trial would have needed longer follow-up for this end-point and was set to detect hip fracture after 20 years, following 10 years' treatment; vertebral fractures were to be assessed from a sample after 5 and 10 years, and the trial was set to detect a halving of risk with 80% power and 5% significance. European and Australian collaboration was necessary to achieve power to detect the effects of HRT on risk of breast cancer.

### **Evaluation: assumptions for exemplar outcomes**

For the *ex ante* assessment of the likely value of the trial, three exemplar outcomes were considered: (1) a 'positive' outcome showing clear long-term benefits related to cardiovascular events and fractures, outweighing any increase in risk of breast cancer, and adding to the benefits of reduced osteoporosis and short-term relief from menopausal symptoms; (3) a 'negative' outcome where there are negative long-term 'benefits' from the therapy, although still short-term relief from menopausal symptoms. The assumptions for the negative outcome were that the cardiovascular benefits were a quarter and fracture benefits a half those for the positive outcome, whereas the increase in risk of breast cancer would be twice that under the 'positive' outcome. This was considered a 'worst case' outcome on current knowledge, i.e. a small cardiovascular benefit suggested but counteracted by a higher breast cancer risk. An alternative inconclusive outcome (2) was the other scenario in which the cardiovascular and fracture benefits were as in the negative outcome, but with breast cancer risk as in the positive outcome. This could present as an 'inconclusive', non-significant result, possibly due to insufficient power.

### **Assumptions for policy implementation scenarios and costs and benefits**

Use of HRT following publication of the trial outcome was predicted to 2027 (20 years) for each of the three exemplar outcomes. These were based on published data<sup>45,46</sup> and modelling current trends in the UK<sup>47</sup> and USA, and on what seemed 'reasonable'. Expert opinion other than from the literature was not used. In each case costs and benefits of HRT use were assessed net of the likely use in the absence of the trial, assumed to be a continuation of the trend up to an assumed maximum, which was the current usage in California.<sup>48</sup> In the absence of the trial, benefits would nevertheless vary according to the exemplar outcomes, although these would not be known and would not affect policy or cost of use.

Following a positive outcome (1), it was assumed that the 50% of women aged 50–64 years currently taking HRT short term (including the 10% continuing to long-term therapy) would all continue on to HRT long term. The net discounted present value of this policy change over 20 years, together with the present value of the trial costs, was estimated at £598 million (*Table 2*). The associated net benefit in terms of normal healthy year equivalents gained was based on the

**Table 2** Implications of the exemplar outcomes and related policy change in the proposed HRT trial

	<b>Positive outcome 1</b>	<b>Inconclusive outcome 2</b>	<b>Negative outcome 3</b>
<b>With trial</b>			
Net therapy/health-service cost discounted at 6% to 1994 (£ m)	1049.70	498.67	223.00
Trial cost (discounted to 1994) (£ m)	47.09	47.09	47.09
Benefits (QALYs million)	4.50	1.96	1.96
<b>Without trial</b>			
Net therapy/health-service cost discounted at 6% to 1994 (£ m)	498.67	498.67	498.67
Benefits (QALYs million)	2.81	1.96	1.11
<b>Net trial implications</b>			
Net costs (£ m)	598.12	47.09	-228.58
Net benefits (QALYs million, not discounted)	1.69	0	0.85
Net benefits (QALYs million, discounted)	0.35	0	0.18
<b>Costs per QALY (not discounted)</b>	£354	Costs for no benefits	Benefits + cost saving
<b>Costs per QALY (discounted)</b>	£1709		Benefits + cost savings

assumptions and on published research on the effects of HRT on quality of life, and estimated as 1.7 million quality-adjusted life years (QALYs) (0.35 million discounted), gained at a cost of £354 per QALY (£1709 discounted). For the inconclusive outcome (2), no change in policy or practice was assumed and so no net benefit or net cost of the policy change. This indicates, for example, the lack of return from a trial that is too small to identify significant effects, so there would be the cost of the trial for no information gain. If there were a negative long-term outcome (outcome 3), a policy change to 'no long-term usage' was assumed. This negative implementation would mean a reduction in net health service and therapy costs which would be far greater than the trial costs; there would also be an increase in QALY benefit from avoidance of the long-term disbenefit. This outcome, the health-service 'dream' outcome, would offer both a cost reduction and a benefit increase.

### Expected value of trial (weighted analysis)

The outcome of the trial was obviously not known *a priori* and so different likelihood or weightings of each of the three considered outcomes were used. These were that each outcome was equally likely (weighting one-third each); that outcome 1 was more likely (weighting 50%) and other outcomes 25%; and, less likely, that the negative outcome 3 was more likely (weighting 50%) and other outcomes 25%.

The inconclusive result (outcome 2) would give the costs of the trial for no benefits (infinite costs per QALY).

Weighting the possible outcomes of the trial by these assumed probabilities for each result allowed estimation of the expected net costs and benefits for the trial. Equal weighting resulted in an estimate of £160 per QALY (£770 discounted); higher weighting to the negative outcome resulted in an estimate of £55 per QALY (£260 discounted). On current knowledge, however, the most realistic assumption was represented by the high weighting to the positive outcome, which resulted in expected health gains in QALYs at a cost of about £240 per QALY (£1150 discounted). This best *a priori* estimate compared with an estimated *ex post* benefit of £350 per QALY (£1700 discounted) for a positive outcome, a health benefit and a cost saving for the negative outcome, and cost for no benefit for the inconclusive outcome.

### Conclusion

The application of the model to the returns to the trial suggested that it was likely to result in a health gain in QALYs at an expected cost of £240 per QALY (£1150 discounted), with a range of -£270 to +£350 (-£1270 to +£1700 discounted), depending on whether the actual result was negative or positive. Comparison with costs per QALY from other studies showed that this represented a lower expected cost than that of

many common health-service treatments, such as that of renal transplant at £2000 per QALY (£4500 discounted) or breast cancer screening at £3000 per QALY (£5000 discounted).

The analysis depended on making assumptions about the long-term effects of HRT which were uncertain and which the trial was designed to clarify. However, given the current knowledge, it was judged that the analysis indicated that the

trial would be a 'good buy' in terms of common health-service practice, whether the outcome was to extend treatment or, if HRT proved to be non-beneficial, to limit its use. In addition, as in any assessment, there was likely to be additional knowledge; in this case it would have given profiles of health states and health-service use through middle and into old age, for a sizeable cohort of women taking HRT and those not taking it. The project was funded.



# Chapter 3

## The synthesised model

### **Development**

In the light of the analysis of existing models, this section sets out a 'synthesised' methodology to address the various problems previously identified, while incorporating the strengths of existing models identified in the literature review. It essentially retains the separation of a first stage that assesses whether and what type of research on a technology is likely to be cost-effective and a second stage that focuses on whether a particular research proposal (for that topic) is likely to be cost-effective, or which of a number of alternative proposals is likely to be most cost-effective. The general model is represented in *Figure 3*. Separate quantitative analyses are applicable to the top and lower halves of the model, assessing whether uncertainty exists and if so in which area, as well as to the stage of assessment of a specific health technology assessment proposal. Before either assessment, data are needed to decide the level of uncertainty about the cost and/or effectiveness of an intervention, and whether this uncertainty reasonably prevents informed adoption or rejection of the technology. If a reduction in uncertainty would, given a particular outcome and the necessary resources, lead to a beneficial change in practice or policy, then a trial should be considered. For specific research proposals, the expected incremental cost-effectiveness ratio of the research proposals could be determined using equation (2). This ratio would be subject to sensitivity analysis to allow for the uncertainty and subjectivity around the inputs to the model. The results of the application of the model would be presented in a report to a research-funding organisation, in which this expected incremental cost-effectiveness ratio (EICER) would be weighed with other payback considerations in the final decision about whether the research project should be funded or not.

The quantitative analyses in the model, as summarised in equation (2), are designed to be applicable at both the technology level (top of the model) and the end of the algorithm, namely whether a specific research proposal is likely to be cost-effective. The analysis is based on three alternative potential broad outcomes to an assessment, and could be extended to include

multiple outcomes as appropriate to the context of a specific technology. The three broad potential outcomes include a favourable conclusion, an unfavourable conclusion and an inconclusive outcome. Each of these conclusions will have an associated level of benefit (or disbenefit) and a level of costs depending on the outcome and effect on dissemination and implementation; there will be an associated probability that the assessment would result in that conclusion. The EICER formula is:

$$\text{EICER} = \frac{C_T + p_1(C_1 - C_c) + p_2(C_2 - C_c) + p_3(C_3 - C_c)}{p_1(B_1 - B_{c_1}) + p_2(B_2 - B_{c_2}) + p_3(B_3 - B_{c_3})} \quad (2)$$

where  $C_T$  = cost of the trial,  $n$  = trial outcome identifier (1 = favourable to intervention, 2 = unfavourable to intervention, 3 = inconclusive result),  $C_n$  = cost associated with likely practice, given trial outcome  $n$ ,  $C_c$  = cost associated with likely practice if trial did not occur,  $B_n$  = likely benefits associated with likely practice given trial outcome  $n$ ,  $B_{C_n}$  = likely benefits associated with likely practice if trial did not occur, given the possible but unobserved outcome of the trial,  $p_n$  = probability that trial will produce outcome  $n$ .

*Figure 3* sets out the full PATHS model, from the generation of ideas for assessment incorporating horizon scanning through to the rejection or funding of a specific research proposal.

The EICER could be used as a tool for health technology assessment funding bodies at the technology level or to assess potential returns from proposed research. It would not be likely that funds were available for all research with a favourable EICER, but by comparing the expected cost-effectiveness ratio of one or a series of research technology areas or proposals within the same area and the cost and effects of continuing with the current provision, a funding body would be in a position to allocate limited research funds to provide the most efficient returns to research. Each assessment would compare a range of alternative implementation patterns following the assessment's conclusion, against the likely practice in its absence, using a probabilistic framework. It

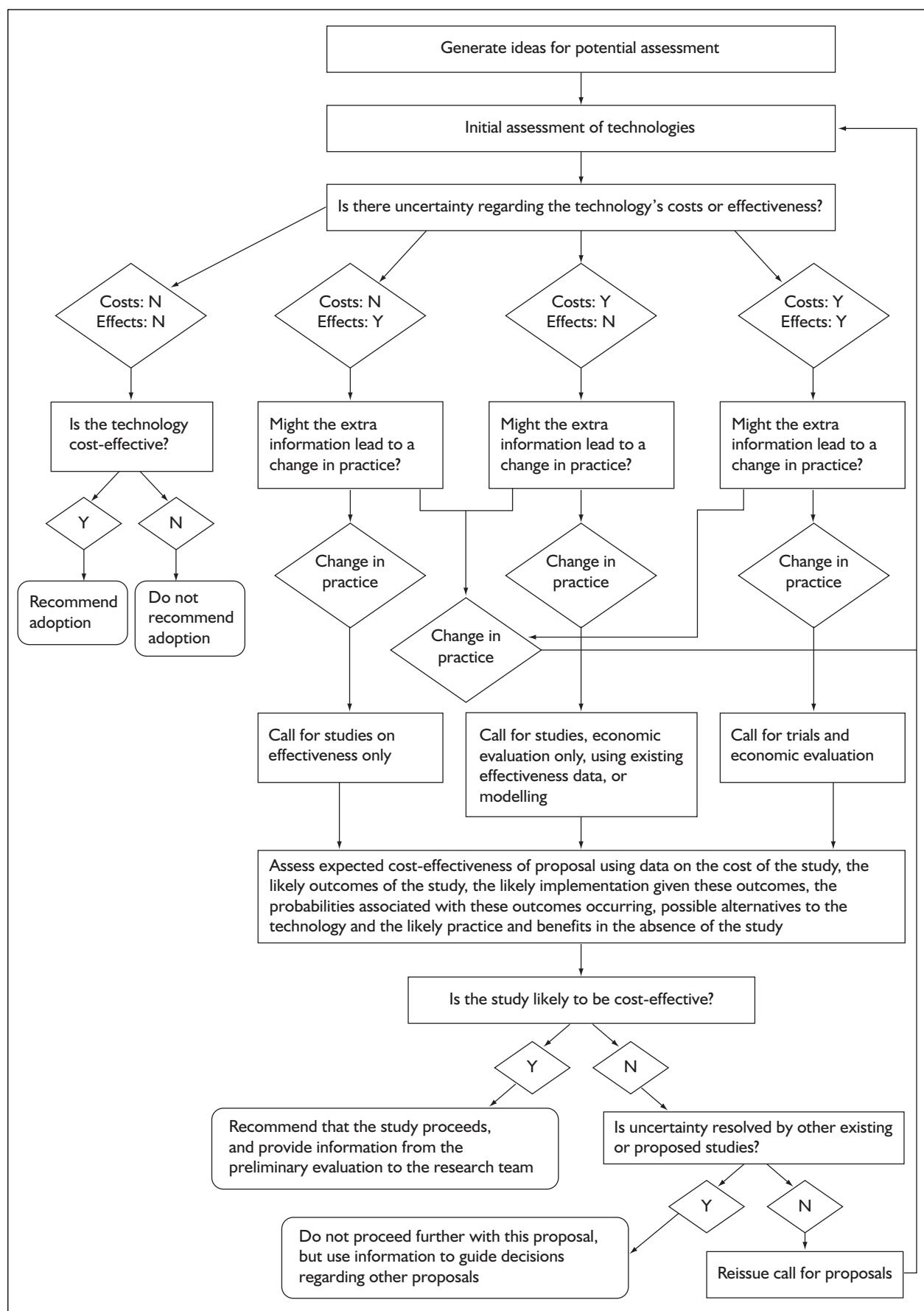


FIGURE 3 General outline of the PATHS model

would measure the effectiveness of assessment in terms of the primary outcome measure of the research proposal itself using existing published data, or expert opinion if necessary.

It would not require major data collection, would be relatively straightforward to operationalise and would be relatively transparent. Alternative values for the net costs, benefits and probabilities for each scenario could be incorporated. The extent of the implementation can be adjusted allowing the evaluation to reflect the likely impact on practice within the health sector, as well as the reduction in uncertainty.

### **Data for the model**

Where empirical data are not available from the literature or routine data sources for use within the model, expert opinion would be used to provide estimates. The choice of experts would depend on the nature of the proposed trial under analysis, in particular the level of intervention, but would represent clinical, health-economic and purchaser expertise to triangulate the estimates. The clinician is important for making predictions in the light of expert knowledge of both the technicalities or intricacies of the subject and current practice, and the likely impact of alternative research findings on decision-making and policy at the clinical and professional level. The purchaser's contribution would be from the view of a more macro decision-maker who has the experience of responding to new evidence and allocating resources between different services. The economist's contribution would be technical (understanding of cost utility analysis and decision analysis) and would usually be made by a researcher who has worked in the field under consideration and so would be cognisant with the issues.

To operationalise the model the following information would be needed:

- the likely development of policy or practice in the absence of the assessment (i.e. current use and likely trends): the counterfactual
- the likely implementation of the experimental intervention following the reportings of the assessment results, given alternative scenarios for the results:
  - the extent to which the technology or service might be adopted
  - how existing policy might affect the level of implementation
  - whether the technology or service would be introduced for all potential candidates, or whether it would be targeted at specific groups

- how other relevant services might be affected following the adoption or change in practice of the experimental intervention
- what measure would be most appropriate or in the case of assessment at project level, whether the outcomes used within the trial are appropriate, whether they are likely to influence future policy and/or practice and whether there are alternative outcome measures that might be more appropriate
- threshold levels of costs, effectiveness or cost-effectiveness that might trigger adoption of the experimental intervention
- probabilities of alternative outcomes of the assessment (positive, negative or inconclusive).

### **Conclusions**

This chapter and the previous one have considered characteristics that a quantitative model for the prioritisation of health technology assessment projects needs to include. There are inevitably trade-offs among the completeness of a model, the resource implications of the modelling exercise and the ease of operationalising the model. A balance has to be struck among them. The model presented here does not aim to replace existing processes of research prioritisation, but would be complementary to the process of peer review and criteria-based approaches that consider benefits that are not directly quantifiable. It aims to fit within the existing processes and to be used within an iterative framework.

Transparency is an essential characteristic of the process as potential applicants need to know how their proposal is to be assessed, and this decision-analytical model is far more transparent than a model using abstract methods to determine the benefits of the research.

Use of a range of expert opinion would help to obviate bias, and sensitivity analysis is relatively straightforward in models based on decision-analytical theory, as the costs, benefits and probabilities of each arm can be easily adjusted. Models in which the benefits are difficult to disentangle may not be easily adjusted by sensitivity analysis to account for uncertainty. An attempt has been made to balance the issues of rigour, transparency and resource requirements.

### **Testing the model**

It was considered important to test how the model might work in practice. Ideally, this would have been done prospectively on a cohort of proposals

in parallel with the routine process to determine whether they would be funded. For those funded, the ‘predictions’ from the *ex ante* evaluation would be compared with the outcomes of the technology assessments and the impact these had on policy and practice. However, such real-time testing would have been excessively lengthy and impossible to undertake within the desired time span of this study. As a realistic approximation, therefore, the authors proposed to test and evaluate the model on four case studies of already funded randomised trials, identified in liaison with the UK NHS R&D HTA programme and the MRC. These case studies would continue to run concurrently with the preliminary modelling exercise then, later, would be completed before the release of results from the trials, so that the evaluation results could be compared with the model predictions using expert professional opinion as to what impact the actual results were likely to have on the future use of the technology.

A further intention was to provide a practical guide to health technology assessment organisations, particularly on the quantitative assessment, that could be used as an adjunct to current assessment procedures to prioritise research or to appraise specific research proposals.

In the original proposal it was planned to carry out eight case studies. In the commissioning process the study was scaled down and four were agreed. The MRC and the NHS HTA R&D programme provided details of their ongoing studies and four from only eight were identified that fitted the criteria of being an RCT, including an economic analysis and reporting within the window of the second year of the study. The last criterion was necessary so that the *ex ante* analysis could be carried out before the trial reported, and to conduct the *ex post* analysis during the study period.

Two MRC- and two HTA-funded studies were selected to include both surgery or other invasive procedures and other non-invasive health services research projects. In the event, only one of the four cases reported on time. One reported late, one had not been reported officially at the time of the analysis, although the researchers have given us confidential access to their results, and the fourth case had been given an extension by the MRC and had not been completed.

The following reports present assessments of these case studies using the PATHS model. The modelling exercises analyse the expected benefits from undertaking each research project, and relate these to the likely costs associated with implementation of results and of undertaking the trial itself.

For each case a summary of the research project is first presented, followed by a report of the interviews and predictions from the expert group, the results from combining the routine data with the information acquired from the expert groups into the model, a discussion of the implications of the analysis and the conclusions. This first section of each case report contains the material that could be presented to a potential funding body to help to inform the funding decision. The final section of each case report presents the actual results and the costs and benefits of likely implementation based on expert opinion and so constitutes the tests of the model. The documents sent to members of the expert group are given in the appendices. The first set of documents sent included project summary, protocols, some relevant references and an outline of the questions asked. The later documents sent after project results were available included the article or report in which the results were published (summary only given here) and questions asked about likely implementation.

## Chapter 4

# Case study 1: the costs and benefits of postnatal midwifery support

### Project summary

The trial of 'The costs and benefits of postnatal midwifery support' aimed to evaluate the effectiveness and cost-effectiveness of a postnatal support service provided to mothers by community midwifery support workers (CMSWs). The CMSW was to provide the following services:

- establish a supportive relationship
- provide a personalised service with regular contact to facilitate the recognition of signs of problems
- reinforce midwifery advice and information about breast-feeding and other issues
- carry out light housework
- encourage women to use the services of healthcare professionals and other agencies appropriately.

The service would be in addition to the standard current midwifery visits.

The trial planned to recruit 720 women over a 6 month period. Half of the mothers recruited were to be randomly allocated to receive up to ten daily weekday visits offered from the CMSW in addition to standard care. The other mothers were to form the control group and receive the current standard midwifery postnatal care programme of ten midwife visits up to the 28th day following delivery and 24 hour telephone access to the midwifery service until the 28th day.

The trial aimed to evaluate whether or not the intervention group receiving the CMSW service had higher health status after 6 weeks relative to the controls, as measured by the general health perception dimension of the Short Form 36 (SF-36). Secondary outcomes of postnatal depression, breast-feeding rates, uptake of benefits, level of economic activity, mother's feelings for her baby and confidence were also to be measured. There was to be an economic analysis to assess the cost-effectiveness of the CMSW service by expressing the additional cost incurred by the experimental intervention per unit of health benefits produced, as measured by the

General Health Perception (GHP) dimension of the SF-36.

### Timings of the trial

The funding application was submitted in September 1995 to start in April 1996. Follow-up was completed in April 1999, with expected and actual publication in July 2000. The initial expert interviews were carried out during August 1998 and the final interviews in August 2000.

### Data from experts

Four experts were asked to participate and were interviewed by either JT or GH. They included the Chief Executive of a Community Trust for the perspective of funding the service change within the perspective of purchasing of community services generally, a Director of Family Health Services for the perspective of detailed professional knowledge of midwifery services and allocating a midwifery budget, a Professor of Midwifery for a wide academic and professional knowledge of both midwifery services and research methodology and outcomes, and a health economist with experience in undertaking economic evaluations of postnatal services for the perspective of cost-effectiveness.

Four people were approached who fitted these categories and all agreed to take part in the study as expert advisors. A pack was sent to each (Appendix 2) including a synopsis of the proposal, copies of four key papers available at the funding decision date<sup>49-52</sup> and the questions on which their opinion was sought.

They were asked how, at the time of funding in 1994, they would have expected the provision of postnatal care to have developed had the trial not been funded, the likely growth of 'selective visiting' or the likely adherence to the national rules of ten daily postnatal midwife visits. They all thought that selective visiting would have been likely to continue to grow in the absence of the

trial, with local midwifery services gradually, officially or unofficially, dropping adherence to the national rules.

They said that if the trial found in favour of the CMSW service then, on average, approximately 25% of midwifery services might well change their practice to incorporate a CMSW service where appropriate. They all thought that in the counterfactual situation, without the trial, such a service might be used in any case, with possibly up to 5% of units introducing such a service within the near future. To what extent and how quickly a service might be introduced after the successful results of a trial were published, would depend partly on the means of dissemination. The ability to implement would depend on the recruitment and retention of staff needed for a comprehensive service. It was suggested by two of the experts that, were the conclusion to be in favour of the CMSW service, midwifery services might accommodate the costs by targeting the service to a specific geographical area, or a set of women identified as being high risk, and that the limited service might be run for, say, 6–12 months as a pilot service to assess its value and acceptance.

Were the results to be negative, that is, the women receiving the CMSW service had a lower health status than those receiving the standard care, there was agreement that adoption of the service would not occur, although all experts said that a few midwifery services might still adopt or retain a limited service despite the conflicting evidence, because it was a new idea, because they were not concerned by cost implications or because they saw advantages in terms of equity and so might use the service for high-risk women only. Were the trial results inconclusive, it was suggested that further research or pilot projects might be undertaken. In this case, it would be likely that the services might continue according to the existing trend without the trial, with a very few services adopting a CMSW service as part of further research or pilot projects, as described above.

### **Outcome measures**

The experts considered the outcome measures to be relevant, although there was considerable scepticism about how meaningful the GHP dimension of the SF-36 would be to policy makers. This was the main end-point and so was crucial. The secondary end-points of postnatal depression and breast-feeding rates were identified as more important and likely to be more appropriate in influencing policy. Several of the experts would like to have seen other outcomes measured,

including the baby's behaviour, family relations, the interaction between mother and infant, and sleeping difficulties. In summary, it was felt that the study could have included more outcomes concerning the baby, the interaction between mother and baby and other family dynamics, and that the main end-point was difficult to interpret.

### **Cost threshold**

The experts were asked whether they thought there would be threshold levels of costs or cost-effectiveness above which the implementation of the more effective protocol would be prevented. All responded that no improvement in the SF-36 would be likely to affect policy unless it were shown to reduce costs, owing to the pressure on existing maternity service resources, the shift of power towards primary care groups and the likely reluctance of general practitioners (GPs) to divert resources from other areas to such a service. However, if the intervention resulted in a significant improvement in breast-feeding rates and/or postnatal depression, this was seen as likely to encourage adoption of the service at an incremental cost of up to £150 per woman. It was thought that improved scores in the GHP of the SF-36 were likely to be highly correlated with improvements in factors such as breast-feeding rates and postnatal depression. An improvement of five points in the SF-36 was not seen as significant for policy, but a 25-point improvement was, and was seen as likely to trigger adoption. Between these two levels of improvement, the interaction of the other variables was thought to be important. In addition, it was thought that the hidden, indirect costs of reorganisation, when added to the direct costs of the CMSW service, might make a small improvement less attractive and unlikely to affect policy. The distribution of any benefits was considered important, with a small but widespread increase seen as preferable to a large but concentrated improvement, if the decision were to provide a blanket service.

The experts all emphasised that there was little slack in midwifery services and that any extra costs would need to be met by cuts from other areas such as antenatal screening or counselling services. An alternative view from one expert was that £50,000 per service could be found if wished from vacancies not filled. It was argued that NHS Trusts would be unlikely to make any significant reduction in acute services for fear of litigation.

These opinions were used in the model to assess the EICER of this trial, outlined in the next section.

## Model parameters, results and discussion

The costs and benefits of the service for a period of 5 years following publication of the trial results were estimated and discounted to present value at the time of trial funding. Five years was considered an appropriate time horizon owing to the rapid change in delivery of midwifery services, which might make results irrelevant after this time.

The service was costed on a per-mother basis. The trial proposal gave the starting salary of a CMSW as £10,375 and the training costs as £70 per CMSW. Based on the protocol information that a CMSW could see three mothers a day and each mother would receive ten visits in the postpartum period, a CMSW would on average see some 66 mothers per annum. Dividing the annual direct cost of providing one CMSW by the number of mothers seen by the CMSW, we arrive at an approximate cost of £157 per mother, which is rounded up to £160 and will be subject to sensitivity analysis. This figure, related to the proportion of mothers expected to receive the CMSW service following the publication of the trial's results, determines the main cost associated with implementation. The proposal defined effectiveness of the trial in terms of the primary end-point of GHP profile of the SF-36. The total population health improvement (or reduction) from the service was therefore identified by multiplying the number of mothers likely to receive the service by the average expected health gains (or losses) in terms of points increase on the GHP.

The expected changes in cost and effectiveness were used to assess the EICER of the trial, expressed as the expected cost per point improvement on the GHP. Scenarios, representing positive, negative and inconclusive outcomes of the trial, are modelled using varying degrees of implementation based on expert judgement and the likely average health benefit.

The birth rate was assumed to be stable over the period concerned (5 years) and it was assumed that the staff dropout rate does not require the training of additional CMSWs.

As reported above, there was agreement that even in the absence of the trial, a small number of midwifery services would introduce a limited midwifery support service similar to the one under evaluation in this study. The model incorporates this prediction in calculating the net costs and net

benefits of the trial: it was estimated that this base level was approximately 5% of the total number of maternities. Any change would be phased in over 3 years.

*Table 3* sets out three possible alternative trial outcomes and subsequent implementation scenarios identified from the interviews. The costs of the trial and the costs and the benefits of the implementation following each trial outcome are calculated for these three scenarios. In each case we have identified the costs and benefits that would have occurred without the trial, assuming the baseline service to 5% mothers indicated by the experts, in order to calculate the net implications of the trial. *Table 4* presents the costs and benefits with and without the trial for each scenario, the net costs and net benefits of the trial, and the incremental cost per point improvement in the GHP. Costs are discounted at 6% and benefits at 2% to 1994, when the funding decision was made.

A positive outcome of the trial, scenario A, results in net discounted costs of £37 million over 5 years. These net costs are offset by an increase in benefits over the counterfactual to the trial of nine million points improvement in the GHP of the SF-36 over the whole population. This implies a cost per point of £4.20. Scenario B is the inconclusive result. It assumes that there are no health gains and there is a limited degree of implementation. As such, the net costs of the trial would be £0.22 million, for no net health gain. Scenario C, the negative outcome, would result in a net saving of £9.2 million owing to reduction in the service, and a small net benefit of about a third of a million SF-36 points improvement, owing to the service being withdrawn from mothers who would have received the service if the trial had not gone ahead. There is no trade-off between cost and benefit of the trial under this scenario; that is, there would be both a cost reduction and a health gain: the dream scenario. Only one of these scenarios will transpire, and so the estimated net outcomes need to be combined in a probabilistic framework to bring together the overall expected value of the trial (in subsequent cases, the experts were asked to estimate the probabilities of each outcome). By assigning probabilities to each of the scenarios, an expected cost per point improvement could be calculated. Three different weightings were used to cover a wide range of possibilities. Combination 1 is an optimistic combination, with a 50% likelihood of scenario A, and 25% of scenarios B and C; combination 2 is a pessimistic combination, with a 50% likelihood of

**TABLE 3** Scenario assumptions

Trial outcome	Implementation scenario
<b>A</b> (positive) Trial shows that CMSW service is effective and cost-effective: <ul style="list-style-type: none"><li>• on average, a 25-point improvement in the GHP profile of the SF-36</li></ul>	<b>A</b> Following the trial, implementation extends to the majority of services, but to a limited proportion of mothers: <ul style="list-style-type: none"><li>• service provided to 25% of mothers in the postnatal period phased in over 3 years</li><li>• no increase in the number of GP visits</li></ul>
<b>B</b> (inconclusive) Trial finds marginal improvement in health status due to CMSW service: <ul style="list-style-type: none"><li>• on average, no improvement in the GHP profile of the SF-36</li></ul>	<b>B</b> Following the trial, the majority of midwifery services do not adopt the service, although several continue to provide the service, and some innovative departments develop a limited CMSW service: <ul style="list-style-type: none"><li>• service provided to 5% of mothers in postnatal period</li><li>• on average, an increase of one GP visit in 10% of the women receiving the CMSW service</li></ul>
<b>C</b> (negative) Trial finds that mothers without the CMSW service have higher health status: <ul style="list-style-type: none"><li>• on average, a five-point reduction in the GHP profile of the SF-36 for those with the CMSW service</li></ul>	<b>C</b> Following the trial, owing to either poor dissemination or the inability to extrapolate the trial results to local settings, a minority of departments continue to provide a CMSW service: <ul style="list-style-type: none"><li>• service provided to 2.5% of mothers in postnatal period</li><li>• on average, an increase of two GP visits in 10% of the women receiving the CMSW service</li></ul>

**TABLE 4** Cost and benefit calculations for the UK

	Scenario A	Scenario B	Scenario C
Trial result:	Positive	Inconclusive	Negative
Without trial			
Net cost (£ m)	18.9	18.9	18.9
Net benefits (million points of GHP of SF-36)	3.20	0	-0.64
Trial			
Trial cost (£ m)	0.223	0.223	0.223
Following trial (5 years)			
Cost (£ m)	55.6	18.9	9.5
Benefits (million points of GHP of SF-36)	12.04	0	-0.32
Net trial implications			
Net costs (£ m)	36.9	0.22	-9.2
Net benefits (million points of GHP of SF-36)	8.8	0	0.32
Cost/point of GHP of SF-36, as a result of trial and subsequent implementation	£4.20	NA	(savings + benefit)
NA: not applicable.			

**TABLE 5** Expected costs and benefits

	Combination 1	Combination 2	Combination 3
Expected costs (£ m)	£16.205	£4.68	£9.307
Expected benefits (million points)	4.48	2.36	3.04
Expected cost/point	£3.62	£1.98	£3.06

scenario C, and 25% of scenarios A and B; combination 3 is the neutral combination, with 33% likelihood being assigned to all three scenarios. The final weighted results are summarised in *Table 5*.

Owing to the high costs associated with scenario A, the expected cost per point is not highly sensitive to the probabilities, and improvement under all three combinations is within the range of £2.00 to £3.60 per point improvement (calculations for *Table 5* use cost and benefits figures from *Table 4* rounded up to the nearest thousand).

## Conclusions

In evaluating the likely returns to the proposed research, the authors have tried to assess the likely payback in terms of discounted future costs and benefits flowing from the implementation of the results, and have also addressed issues around the research design, dissemination and likely implementation of the research findings.

The main conclusion from this study is that the expected cost per one-point improvement in the GHP scale of the SF-36 is between £2.00 and £3.50. When this is translated into the gain suggested by the expert panel as one which would bring about change, that is a 25-point improvement, it implies a maximum net cost of introducing the technology of approximately £50 to £90 per mother. This cost would be well below the experts' threshold of £150 per mother for implementation were there a significant improvement in breast-feeding rates or depression, and which they thought would be highly correlated with a 25-point improvement in the GHP.

The proposal is therefore assessed as 'good value for money' on the basis of the estimated stream of benefits and costs compared with the threshold level for SF-36 points and costs given by the experts. An important issue arises here in the translation from 'statistically significant' to 'clinically important'. The experts were not clear that the reporting of an improvement of 25 points in an SF-36 scale would bring about a change in practice, owing to the lack of understanding of its meaning by decision-makers. The expert group was more interested in the secondary outcome measures, especially breast-feeding rates and postnatal depression, as being more likely to influence commissioners and providers of family and maternity services, although they thought that

the primary end-point was likely to be correlated with these.

According to the experts, in the face of 'inconclusive' evidence, the net changes in policy compared with what would have happened if the trial were not funded would be minimal and therefore the net costs and benefits, although very low, would bring no returns to the cost of the trial. In contrast, scenario C, were it to be the outcome, would be an ideal scenario and the best value for money, with a net saving over 5 years of £9.2 million, and a net increase in health status of 0.32 million points in the GHP of the SF-36. Neither of the two conclusive outcomes is therefore unattractive.

The overall conclusion would be a recommendation to fund the trial, but with two design changes. If the results are to convince providers of maternity services, it would be preferable that the trial was set to detect a significant change in breast-feeding rates or in postnatal depression (the trial includes but is not designed to detect any such differences), with emphasis on these rather than the observed differences in the GHP profile of the SF-36. The experts would also like to see other child-related outcomes, as set out above. The other issue that was emphasised would be to consider subgroups of women at relatively high risk, as a service directed at these women might be more cost-effective and more likely to be implemented.

This report was written following interviews with the expert advisors in July 1998, 2 years before the publication of the results of the trial, which follow.

## Actual results of the trial

The results of the trial were published in March 2000 as a *Health Technology Assessment*.<sup>53</sup>

The main results were as follows.

- In total, 623 women were randomised, they were well matched between intervention and control groups and there was a good response to follow-up.
- At 6 weeks there was no evidence of a significant difference between the groups for the primary outcome (GHP domain of the SF-36).
- There was a non-significant trend for the control group to have better Duke Functional

- Social Support (DUFSS) score and Edinburgh Post Natal Depression (EPDS) score at 6 weeks.
- There was no significant difference in breast-feeding at 6 weeks.
  - At 6 months both groups had similar health status.
  - Satisfaction with the new service was the highest of all postnatal services received.
  - There was no difference between the groups in the use of other services (GP contacts, hospital services, prescriptions or medicines for mother or baby).
  - The total mean NHS costs to 6 month follow-up were £180 higher for the intervention group (confidence interval £79.60 to £272.40). This included setting up and running the service.

The researchers concluded that although women valued the service there was no evidence of any health benefit at 6 weeks or at 6 months and no difference in health-service use, but an extra cost of £180 per woman, and that additional studies are needed to identify support-related outcomes of importance to postnatal women and to compare the effectiveness of different models of antenatal and postnatal support.

## **Testing of the model**

Following the publication of the report, a copy was sent to each of the experts, together with a second set of questions (Appendix 3).

Related to the three possible scenarios considered, the outcome falls between inconclusive and negative and would fail the experts' cost threshold for implementation.

## ***Ex post interviews with experts: their predictions of effects of the trial results on subsequent policy***

### ***Implementation***

All of the experts said that they would not recommend nor expect take-up of the CMSW service, and that where in use they would recommend and expect reduction of the service. However, they all also said that they thought the service might continue for certain at-risk groups.

Each expert suggested a different at-risk group: one suggested ethnic minority groups, particularly Asian women, who tend to have low breast-feeding rates, and another suggested low-income families. (Some of these may have been specifically excluded from the trial in the exclusion criteria of not speaking or understanding English, having a baby requiring special care in a baby unit, or being

aged 16 or below.) If introduced for an at-risk group, they thought this would happen over 3 years. If reduced, this would take 1 year and would require careful handling and presentation of the evidence. It was anticipated that there would still be a minimal service operated across the country, possibly to some 3% of mothers (from an existing base level of possibly 5%). One of the experts advised that there should be continued service for at-risk groups, with audit and further research including willingness to pay.

### ***Outcomes***

The experts were asked which of the outcome measures influenced their predictions and to put the measures in order of importance to them. Three put breast-feeding and depression first and one said that these were the only outcomes of relevance. One put patient satisfaction first. Three added SF-36, SF-36 GHP (the major end-point used for power calculation) as third or fourth; client satisfaction (the outcome that favoured the intervention) was dismissed by one, put in first place by another and in fifth by the other two, so there was disagreement about its role.

### ***Dissemination***

To achieve dissemination of the results it was suggested that they should be published in a midwifery journal and sent to heads of midwifery departments (expert A); then they should be disseminated to midwifery liaison committees, press releases, Community Health Councils, midwives and support workers and there should be feedback to The Netherlands, where the service is in general use (expert B); and funds should be sought for a national meeting on the current state of postnatal support possibly linked to Sure Start and other interventions (expert C). The fourth expert also recommended peer-reviewed publications and conferences (expert D).

### ***Overall comments on the trial***

- Expert A thought that, in general, postnatal visits should be reduced and early discharge extended, and the savings spent on teaching parenting skills and giving vouchers for takeaways where postnatal domestic support was needed.
- Expert B thought that the study was very useful as it provided proper evidence-based results, and that further research should look at substituting midwives and/or health visitors for CMSWs.
- Expert C thought that it was a good study, carefully conducted with appropriate methods, and was being scrutinised by researchers

- evaluating other voluntary and health visitor social support.
- Expert D thought it a useful, well-conducted study, but limited by not using sensitive enough tools, long enough follow-up or at-risk groups.

## Modelling *ex post* predicted outcomes

The *ex post* expert agreement was for a reduction in the service (which was currently thought to be running at a level of implementation of about 5%) to a level of about 3%, and that this reduction would take about 1 year to implement.

This would bring about a cost saving of the service to 2% of all mothers, saving £180 per mother. As a 5 year time horizon is used for this case, the saving would be for years 2–5 discounted at 6% back to the funding of the trial (1994), which represents a saving of £5.86 million. The cost of the trial was £0.223 million, leaving a saving of about £5.6 million. There could be some loss of mothers' satisfaction from the service, but also some reduction in depression, as lower rates of depression were noted in the control group. The savings would be higher if a longer time horizon were considered.

The *ex post* results confirm the value of the study as assessed in the *ex ante* evaluation; the negative outcome, resulting in a reduction of the service, gains the maximum value from the study. The results are dependent on the assessment of the likely reduction in use of the service and on the baseline estimates given by the experts. It is interesting, therefore, to consider the sensitivity level for the break-even point of service reduction. The 5 year savings would match the cost of the trial if the reduction in service were considerably lower than 2%, i.e. only 0.08%. The results are therefore robust and the model successfully predicted the trial to be cost-effective.

This is an area in which there has been little research. It is therefore likely to have the added benefit of training the researchers employed on the trial and the participating services. It may also be considered useful in suggesting future research, particularly in developing or identifying appropriate outcome measures in an area of considerable public health concern and sensitivity.



## Chapter 5

# Case study 2: an RCT of infusion protocols in adult pre-hospital care

### Introduction

The NHS R&D HTA-funded trial 'A randomised controlled trial of prehospital intravenous fluid replacement therapy in serious trauma'<sup>63</sup> is the second case study for testing out the model. The aim of the modelling exercise is again to analyse *ex ante* (as at before funding) the expected returns from undertaking the research project, and to relate these expected returns to the likely returns estimated *ex post* (after publication of results), associated with subsequent implementation and undertaking of the trial itself.

### Project summary

The project aimed to evaluate the use of cannulation and fluid infusion by ambulance personnel in severely injured adults, in response to the debate among professionals and within the published literature about the uncertainty surrounding the benefits of intravenous fluid replacement on-scene. Evidence surrounding the effectiveness of on-site infusion was inconclusive. Some argued that on-site infusion stabilised patients haemodynamically, resulting in a better patient outcome. Others argued that infusion, rather than stabilising the patient, might result in elevation of the blood pressure, increased blood loss and delay associated with establishing an intravenous line. This might lead to an increase in the time taken to transport the patient to the hospital, resulting in a worse outcome. These two opposing views were tested, comparing two protocols in the trial as follows.

- **Scoop and run:** following basic life support, the patient was to be transported immediately to hospital if the time from arrival on-scene to arrival at the hospital was estimated to be over 40 minutes.
- **Field stabilisation:** following basic life support, an intravenous line was to be established and the patient infused with a crystalloid/colloid mixture, unless the estimated time from arrival on-scene to arrival at the hospital was less than 15 minutes.

The unit of randomisation was the paramedic: 420 paramedics in two ambulance services were randomised to follow one or other of the two protocols under strictly defined circumstances. Halfway through the study the randomisation was crossed over to avoid bias.

The trial aimed to determine which protocol was the more effective primarily in reducing mortality but also morbidity over a 6 month period, and which was the more cost-effective from an NHS perspective. To investigate this, the study followed up patients over a 6 month period and recorded all contacts with health services. General health status was measured by the SF-36. The trial economic analysis was to be expressed by the marginal cost of the field stabilisation protocol service alongside any marginal benefits in terms of reduced mortality and/or disability.

### Timings of the trial

The funding application was submitted in July 1994 to start in September 1995. Follow-up was completed in September 1998; the report was made available to the authors in June 2000 and published in November 2000. The first set of interviews was carried out in February 1999 and the final interviews in July 2000.

### Data from experts

Four experts were approached to participate and were interviewed by either JT or GH. The panel included a Consultant in Emergency Medicine for professional knowledge of the procedures and outcomes, a Medical Director of an Ambulance Service for professional knowledge of the service, paramedics and policy, a Chief Executive of an NHS Trust for the broad expenditure priorities, and a health economist with experience in undertaking economic evaluations of pre-hospital services for the perspective on cost-effectiveness. All those approached agreed to take part and a pack was sent to each (Appendix 4), including a

synopsis of the proposal, copies of eight key papers<sup>54-61</sup> (all short) available at the time of the funding decision data, and the questions on which their opinions were sought. They were asked how, at the time of the trial funding in 1995, they would have expected the provision of pre-hospital care to have developed had the trial not been funded, in particular whether or not, or what proportion of patients would be cannulated before transfer to hospital.

They were in agreement that there would have been a gradual shift away from field stabilisation, which was already losing favour, towards a policy of scoop and run, unless the transfer time implied that stabilisation was necessary. Each said that if the trial found in favour of the scoop and run protocol, this shift away from field stabilisation would be accelerated. If, however, the trial found in favour of field stabilisation (i.e. mortality was significantly lower in those patients cannulated on-scene, and the protocol was cost-effective), then there was likely to be a switch in policy where appropriate. However, the rate at which ambulance services would be likely to switch to employ this protocol would be slower than if the scoop and run protocol was favoured, as it would be counter to the current trend. Were the trial results inconclusive, it was felt that the current trend would continue, but perhaps at a slower rate than if the trial found in favour of scoop and run. These projections were incorporated into a 10 year model for each scenario shown in detail in the next section. There is little definitive research in this area and it was felt that the findings were likely to hold for a relatively long period, although new developments might overtake after this time.

The experts were asked which of all the outcome measures used within the study was the most likely to influence practice among ambulance services. The accident and emergency (A&E) consultant said that the quality of life measures would be most important, as mortality would be most unlikely to be affected by the protocol difference. The three other experts agreed on mortality, the primary outcome measure of the trial, as the most important outcome. In accordance with the trial protocol, mortality was used to calculate the EICER in this exercise, while recognising that morbidity and disability are important outcomes not considered explicitly within the quantitative modelling.

The experts were asked whether they thought there were threshold levels of costs or cost-effectiveness above which the implementation of

whichever protocol was found to be more effective, would nevertheless not be used. They unanimously agreed that cost differences between the protocols to the NHS were likely to be small and not likely to be barriers to implementation. Implementing changes that went against intuition or the existing trend might be difficult, however, but not for cost reasons. They considered that organisational or professional barriers to the adoption of the more effective protocol were more likely than cost barriers.

They were also asked, given their knowledge of current evidence, what they thought the results of the trial were likely to be. The general agreement was for a result in favour of the scoop and run protocol as the most likely outcome, with a probability ranging from 0.4 to 0.5. Probabilities for an inconclusive outcome and for favouring field stabilisation were each put between 0.25 and 0.3. The A&E consultant said ominously that uncertainty is always the real outcome of such studies.

These opinions were used in the model to assess its likely incremental cost-effectiveness ratio.

## **Model parameters, results and discussion**

The costs and benefits flowing from the different scenarios were calculated for a period of 10 years on a per life saved basis, according to the likely change in practice. Differences in patient management as a result of different patient outcomes between the two protocols or the protocol itself would occur on-site, during transport to hospital, in the A&E department, in inpatient departments, and in the primary and community care setting. There could be further cost or quality of life differences if patients recovered more quickly after one protocol, or there were more seriously disabled survivors. The unit cost of each service was identified using the 'Unit costs of health and social care' 1998, as determined by the Personal Social Services Research Unit.<sup>62</sup> Likely levels of demand for these services were suggested by the expert advisors.

Three possible alternative trial outcome scenarios were explored: favouring scoop and run, an inconclusive outcome and favouring field stabilisation. Mortality associated with each protocol, and the net effect on practice following the trial's conclusion were assumed to vary for

**TABLE 6** Proportion of relevant trauma patients treated by scoop and run and field stabilisation protocols in the absence of the trial

	Year									
	1	2	3	4	5	6	7	8	9	10
Scoop and run (%)	50	60	70	70	70	70	70	70	70	70
Field stabilisation (%)	50	40	30	30	30	30	30	30	30	30

**TABLE 7** Scenario assumptions

Trial outcome	Implementation scenario
<b>A</b> Trial reports that scoop and run is more effective: • 10% mortality in patients receiving scoop and run against 20% mortality associated with field stabilisation, resulting in a favourable incremental cost per life saved for scoop and run	<b>A</b> Following the trial, the shift towards scoop and run is accelerated: • 90% of all relevant patients are treated using the scoop and run protocol within 4 years
<b>B</b> Trial does not provide conclusive evidence regarding the effectiveness of field stabilisation compared with scoop and run: • minor improvement in mortality associated with scoop and run (20% vs 20%)	<b>B</b> Following the trial, the existing trend, i.e. the gradual shift towards using scoop and run, occurs, but at a much slower rate: • 70% of relevant patients receive scoop and run by the sixth year after the trial
<b>C</b> Trial reports that field stabilisation is the more effective protocol: • 10% mortality associated with field stabilisation vs 20% mortality associated with scoop and run	<b>C</b> Following the trial, there is a gradual reversal of the existing trend: • 90% of patients are stabilised on-scene by the fifth year

each scenario, based on the trial application and the expert opinion as set out in *Table 7*.

It was assumed that the number of trauma patients per annum would not change significantly over the period concerned. The calculations are based on the mean number of trauma patients treated by a non-metropolitan ambulance service each year.

As reported in the previous section, there was agreement that a gradual policy shift towards scoop and run was currently occurring, and would have been likely to continue in the absence of the trial. *Table 6* shows the assumed pattern of practice over the 10 years in the absence of the trial, based on the expert predictions.

*Table 7* presents the alternative trial scenario outcomes considered and the subsequent likely implementation patterns, based on the expert opinion. These are described in terms of the significant difference in mortality observed in the

trial between the two protocols. To allow for potential variations, calculations were made for each scenario using alternative assumptions. One assumed a relatively high level of health services utilisation. Variations in cost-effectiveness may occur owing to lower levels of mortality resulting in post-trauma survival of people with high levels of disability or morbidity, requiring higher levels of resources (and therefore higher levels of expenditure) in the 6 months following the trauma. *Table 8* presents the assumptions of the unit costs, based on data from the protocol and from 'Unit costs of health and social care' published by the Personal Social Services Research Unit.<sup>62</sup> It also sets out alternative levels of health services utilisation.

For each scenario, the costs of the trial and the costs and the benefits of the implementation following each trial outcome were calculated using the assumptions in *Table 8* based on published data and expert opinion. For each implementation scenario the costs and benefits that would have

**TABLE 8** Unit costs and utilisation rates

	Low utilisation	High utilisation
Ambulance cost (£/minute)	4.46	4.46
Journey time (scoop and run) (minutes)	40	40
Journey time (field stabilisation) (minutes)	60	60
A&E cost (£)	150	150
ITU cost (£/day)	300	300
ITU length of stay (days)	1	2
Ward cost (£/days)	200	200
Ward length of stay (days)	6	8
Readmission cost (£/day)	200	200
Readmission length of stay (days)	4	7
Outpatient cost (£/visit)	80	80
No. of outpatient visits	3	6
GP cost (£/consultation)	15	15
No. of GP visits	2	3

ITU: intensive therapy unit.

**TABLE 9** Cost and benefit calculations: high health services utilisation (UK)

	Scenario A	Scenario B	Scenario C
Trial result:	In favour of scoop and run	Inconclusive	In favour of stabilisation
Without trial			
10 year treatment cost (£ m)	786	729	758.1
Benefits (discounted lives saved)	196,450	181,270	188,750
Trial			
Trial cost (£ m)	0.2603	0.2603	0.2603
Following trial (10 years)			
10 year treatment cost (£ m)	797.1	730	803.9
Benefits (discounted lives saved)	200,300	181,270	199,400
Net trial implications			
Net costs (£ m)	11.4	0.87	46.1
Net benefits (thousand discounted lives saved)	3850	0	10,650
Cost per life saved	£2960	NA	£4330
NA: not applicable.			

**TABLE 10** Cost and benefit calculations: low health services utilisation

	Scenario A	Scenario B	Scenario C
Without trial			
10 year treatment cost (£ m)	515.3	479.1	497.6
Benefits (discounted lives saved)	196,450	181,270	188,750
Trial			
Trial cost (£ m)	0.2603	0.2603	0.2603
Following trial (10 years)			
10 year treatment cost (£ m)	521.3	479.5	529.6
Benefits (discounted lives saved)	200,300	181,270	199,400
Net trial implications			
Net costs (£ m)	6.3	0.26	32.3
Net benefits	3850	0	10,650
Cost per life saved	£1640	NA	£3030
NA: not applicable.			

**TABLE 11** Probabilities of outcome combinations

Trial outcome	Probabilities	
	Expert mean	Equal weights
Favours scoop and run	0.450	0.33
Inconclusive	0.275	0.33
Favours field stabilisation	0.275	0.33

**TABLE 12** Expected incremental costs and benefits from the trial: high utilisation

	Experts' weighting	Equal weighting
Expected costs (£ m)	18.031	19.246
Expected benefits (lives saved)	4660	4830
Expected cost/life saved	3870	3980

**TABLE 13** Expected incremental costs and benefits from the trial: low utilisation

	Experts' weighting	Equal weighting
Expected costs (£ m)	11.880	12.938
Expected benefits (lives saved)	4660	4830
Expected cost/life saved	2550	2680

occurred without the trial have also been estimated and subtracted. This provides the net implications of the trial, presented in *Tables 9* and *10*, together with the net cost per life saved for each scenario, under the high and low utilisation assumptions.

*Tables 9* and *10* present the results of the modelling exercise. The high utilisation group would relate to those with higher disability and morbidity, as they would need more contact with health services in the aftermath of a traumatic accident. Were the results in favour of scoop and run (scenario A), the net costs over 10 years are estimated to be £11.4 million for the high utilisation setting, and £6.3 million for the low utilisation setting. The reduction in mortality associated with the associated policy change would lead to approximately 4000 lives saved over 10 years, at a cost per life saved of £3000 or £1600 for high or low use of services, respectively. Scenario B, the inconclusive outcome, would result in a small increase in costs owing to the cost of the trial and slower trend to field stabilisation after the trial than would have occurred had the trial not been funded, but with no difference in mortality. Scenario C would result in higher net cost levels. This is due to higher costs associated with longer ambulance journeys (see *Table 8*). In addition, the

potential reduction in mortality does not impact as quickly, owing to the slower diffusion rate against the previous trend. The net costs of this scenario after 10 years would be £46.1 million and £32.3 million for the two service utilisation settings. The reduction in mortality would result in 10,650 lives being saved over the 10 years, giving a cost per life saved of £4330 for the higher and £3000 for the lower service use. To obtain the net returns to the trial, the estimates need to be combined probabilistically to determine the EICER. The probabilities assigned were as shown in *Table 11*.

By weighting the net costs and net benefits for each outcome using the expert or equal weight probabilities in *Table 12*, the EICERs of the trial and the subsequent implementation have been calculated for each combination. These are given in *Table 13*.

Turner *et al*<sup>63</sup> showed that the distribution of ages of patients attended by an ambulance is relatively even. Taking the mean age of a patient as 45 years and the normal life expectancy as 75 years, and assuming that the saved patient survives and lives a healthy normal life, the expected incremental cost per life year, of the 30 years saved, under the most pessimistic assumptions would be approximately £130 per life year saved (and

approximately £85 for the lower estimate). The lower estimate assumes that the survivor lives at full health.

## **Conclusions**

These estimates could represent advice to a research funding body about the likely returns to the proposed research. The conclusions relate to the likely payback in terms of costs and benefits and EICERS, and include useful insights about the research design and necessary dissemination, which would be of interest to a research team as well as to potential funders.

Given the assumptions and the expert opinion, a conclusion in favour of either protocol appears to offer significant value for money, as does the overall weighted expected cost per life year. The inconclusive outcome (B) does not provide positive returns. The EICER of the trial would be £2500–£4000 per life saved, or £85–£130 per life year saved. This is low compared with nearly all the current treatment costs and so would be cost-effective. Were the trial to find in favour of the field stabilisation protocol, but the shift away from the current trend towards scoop and run did not occur, or was very slow, then the cost per life saved by the trial would be higher.

The trial, were it to reach either conclusive outcome, would be clearly cost-effective when considering predicted reductions in mortality. The other outcomes measured in the trial may also be important. Should the reduction in post-traumatic incident mortality result in significantly more disabled individuals surviving, or increase the number of individuals requiring psychotherapeutic interventions, then the subsequent utilisation of health services would increase. In these cases the quality of life is likely to be compromised, and the cost per QALY is likely to be higher than the cost per life year saved reported above.

The estimates reinforce the importance of effective dissemination and the implementation of policy supported by high-quality evidence. A number of the experts stressed the importance of bodies such as the Joint Royal College Ambulance Liaison Committee, and the influence that such organisations have on effecting change. However, questions were raised about how quickly ambulance services would introduce change, as this was thought to be slow, and it was thought that paramedics would be reluctant to change their behaviour if the proposed change went

against their own intuition, personal preference or a current trend. The growing importance of clinical governance, nonetheless, should assist this process. The framework and facilities for the effective audit of practice against best evidence should be important factors in effecting change.

With the caveats above, the application of the model to the proposed trial suggests that it offers a good return. This report was written following interviews with experts in July 1998, nearly 2 years before the draft report of the results was available.

## **Actual results of the trial**

These results are based on the draft report submitted to the HTA committee and received in confidence by the PATHS team in June 2000.<sup>63</sup>

The main results were as follows.

- In total, 1309 patients were entered into the study, 699 (53.4%) treated by paramedics operating the fluids protocol A and 618 (46.6%) operating the no-fluids protocol (scoop and run) B. The randomisation worked well, with no significant differences between the treatment arms in incident characteristics, ambulance performance times or patient and injury characteristics, apart from slightly more moderate and severe head injuries in the fluids protocol group A (25.3 versus 20.3%).
- However, protocol compliance by the paramedics was very poor. Only 31% of protocol A patients received pre-hospital fluid (non-compliance 69%). Eighty per cent of protocol B patients were not given fluids (non-compliance 20%); that is, non-compliance overall was about 46%, with only about one-quarter receiving fluids while about three-quarters did not. The estimated odds ratio (OR) of being given pre-hospital fluids when treated by the fluids protocol compared with the non-fluids protocol was only 2.09 [95% confidence interval (CI) 1.53 to 2.81].
- Of the 699 patients in the fluids arm A, 73 died within 6 months (10.4%), and 60 of the 610 patients in the non-fluids arm B died (9.8%), giving a crude OR for deaths when managed by the fluids protocol A of 1.09 (95% CI 0.73 to 1.54). Excluding the 26 patients whose cause of death may not have been trauma related reduces the OR to 1.04 (95% CI 0.69 to 1.55). Excluding 17 patients who may have been dead at the time the ambulance crew arrived reduced the OR to 1.04 (95% CI 0.70 to 1.53).

Adjustment for age, injury severity and whether unconscious at the scene did not significantly alter these odds ratios.

- Health status questionnaires were sent to all 878 surviving patients identified within 7 months of their accident, including an SF-36 questionnaire. The response rate was similar in the two arms (62.9 versus 64.6%). For all eight dimensions of SF-36, protocol A patients (stabilisation) reported a better score, but only one difference was statistically significant and none was at a level considered to be clinically significant.
- For subgroup analysis of each of eight characteristics (ambulance service area, presence of doctor on scene, paramedic contact time, injury severity, emergency surgery, type of injury, type of area, before or after protocol cross-over), there was no evidence of any difference in mortality rates or composite outcomes.
- There were no economically significant cost differences between the two groups.
- The researchers conclude that there is no evidence that the fluid protocols are doing any significant harm in these blunt instrument trauma patients. They say that even though they seem to be associated with an increased risk of death, this could be remediable by altering fluid protocols, although this extrapolates beyond their result. This could be because early fluids do little harm, or because only one-quarter of patients are given them and the protocols did not appear to alter this proportion; that is, the actual treatment was very similar in both arms and so the intention-to-treat analysis is highly confounded, which makes it difficult to draw any conclusions.
- The researchers recommend that ambulance services should concentrate on avoiding unnecessary delays and speed patients to definitive care in hospital.

### **Their recommendations for further research**

- They estimate that patients given fluids spent 12–13 minutes longer on-scene than patients not given fluids, but times were largely unaffected by protocol (although as a result of the poor compliance treatment was not strongly related to protocol either). Scene time may still be the critical issue and needs urgent investigation.
- Further research should compare strict rather than discretionary protocols and only crystalloids should be permitted. Ways of separating fluid infusion and scene time delays should be sought.

- They conclude that the fluids issue remains unresolved, and ask whether there is a similar question for fluids in hospital pre-definitive surgery care: are fluids appropriate in A&E? Do the same questions apply about speed to theatre? Can a blunt trauma trial to prevent fluid resuscitation before theatre arrival be organised?

### **Testing the model**

Following our receipt of the draft report and with the agreement of the authors, a copy was sent to each of the experts, together with a second set of questions (Appendix 5). Related to the three possible scenarios considered, the outcome falls clearly in the inconclusive scenario.

#### ***Ex post expert opinion***

##### ***Change of practice***

The A&E medical director said emphatically that if protocols are to be used strictly then they must be evidence based, disseminated and audited, otherwise they are a waste of time as there will not be compliance. The implication was that this was the reason why the trial protocols were not adhered to, and why the trial does not provide evidence for change. She said that the majority of services now accept that field stabilisation is inappropriate for patients with haemodynamic instability, and the priority is now for airways management, oxygenation, ventilation and ‘go’.

The other experts also said that the results would not change practice and that a very high percentage of services and paramedics use, and should use, scoop and run.

##### ***Timescale of change***

The experts thought this question to be redundant as the results were unlikely to affect practice. They said that practice has already changed, with increased emphasis over the past couple of years on the need for rapid conveyance to hospital once airways had been freed and spinal immobilisation carried out, and that this emphasis was likely to continue and was already the most common practice.

##### ***Protocols***

The experts were asked whether the role of protocols might be questioned outside the range of the trial. One said that they might be, and the other thought that this was not applicable given the results.

### **Dissemination**

Although not appropriate in this instance, the experts thought that the important methods of dissemination would have been via the Joint Royal College Ambulance Liaison Committee, via training (both initial and updating) and protocol usage.

### **Outcomes**

One expert said that mortality was the most important end-point, followed by length of stay in the ITU, SF-36 at 6 months, broad disability and finally SF-36 at 2 weeks. The other expert clinicians thought that mortality was the least likely outcome to be affected (although of course it is important) and thought that length of stay in ITU and the other quality of life issues were more relevant. For two of the experts this was a switch from their previous opinion.

### **Overall comments**

One expert said that there had been much advance publicity about the study and that changes had already been made, and that the issue of protocol compliance is very important, is always assumed but clearly does not happen. The issue of extra on-scene time is important and needs highlighting, as the study appears to confirm that it conveys no benefits. The issue of colloids versus crystalloids is now history, as most services now use crystalloids.

Another expert said that the important thing to come from the study was that it highlights and validates the difficulties of research in this area, and that the study has put down a marker. He saw the problem of the study as that of confounding. First, the trial confounds the role of the paramedic with that of fluid replacement, and putting these two variables together in the trial was bound to result in insignificant results as they are two different effects. He thought that there was also confounding due to lack of stratifying by competence. As paramedics are acting in an emergency situation they will do what they think is best in the circumstances and/or what they can do best. They would feel compelled to intervene if they thought it necessary, but if they were not sufficiently skilled at cannulation they would not wish to take responsibility for it. So, he thought it important that paramedics in such a trial need to be well trained and skilled, taken into confidence and be in agreement with the trial. He also noted the unfortunate and ironic fact that patients who are in the greatest need of fluid replacement, those who have lost most blood, are the most

difficult to intubate, so the patient most in need is least likely to receive fluid replacement. He thought that the results as such were likely to reinforce existing practice and not change anything. He agreed with the authors that a study of fluid replacement needs to be done, that the question is unresolved and is central to predefinitive care, and needs to be evaluated independently from paramedic skills. He made the observation that UK protocols tend to follow US research, which does not capture the unique UK circumstances. Blunt trauma, which is most common in the UK, is less sensitive to intervention than is sharp trauma, which is most common in the USA. In the former, victims tend to die quickly or heal themselves, so trials for blunt trauma need larger numbers than do trials for sharp injuries. He concluded that although this study has not moved the knowledge base on, it has nevertheless moved on the research by identifying its practical problems. Early attempts at research are inevitably less than perfect. He feels that the NHS concentrates too much on service, to the detriment of training and research. He applauded this work for paving the way for future research in the area.

Another expert expressed concern about the trial's evidence of paramedics' poor compliance with protocols.

### **Testing of ex ante predictions**

The results of the trial fall clearly into scenario B (inconclusive) and as such are unlikely to influence current practice. The cost of the trial is not recouped by reduced service costs or added benefit, although there is likely to be some valuable research experience.

All of the experts agreed that the trial had conferred some knowledge benefits, particularly relating to the non-compliance with protocols and the need to investigate this. There was also a strong plea from the A&E consultant expert that the incremental nature of research should be recognised, that methodological mistakes should be acknowledged and that further research in this area should be encouraged. The need to examine the role of fluid replacement in predefinitive care, whether prehospital or in A&E before surgery, was expressed strongly by the two clinicians, and that it should be examined in a situation where it is not confounded with the paramedic's role.

### **Conclusions**

The results were clearly affected by the lack of equipoise, as the evidence of the benefit of

protocol B, scoop and run, has been growing, particularly over the past few years. This was not the only reason for the lack of protocol compliance, however, as there was also not inconsiderable non-compliance of 20% with the scoop and run protocol, which bucked the trend, so a similar proportion of paramedics in both arms carried out scoop and run despite the protocol requirements, probably as they would have done without the protocol. There was some indication that non-compliance was greater in one of the research areas and could be related to entrenched preference, practice or skill; it was also weakly related to case-mix and so to some degree to the relevance to the individual case.

The overall conclusion is that the question of fluid replacement use is one that requires answering by research in the UK setting and that there was

good reason to fund this study area. That the study has not answered the question was due partly to factors outside the control of the researchers or knowledge of the funders: the strong movement towards favouring one protocol, and the paramedics' non-adherence to protocols. These elements were picked up to some extent by the experts at the *ex ante* interviews, although this was after the funding. The results and problems with the execution of the trial emphasise the importance of knowing the baseline practice situation, making sure that key personnel are on board with the study and monitoring for, rather than assuming compliance; non-compliance is a common organisational problem in managing RCTs. The study's problems tend to confirm the strength of the Bayesian approach in health-services research<sup>33</sup> to allow for a background of changing knowledge and practice.



## Chapter 6

# Case study 3: MRC small aneurysm trial of early surgery or observation for small abdominal aortic aneurysms

### Introduction

This chapter presents the third case study using the model. Again, the modelling exercise will analyse the expected benefits from the research project, and relate these to the likely costs associated with likely implementation of results and of undertaking the trial itself.

### Project summary

The project 'Early surgery or observation for small abdominal aortic aneurysms?' known as the MRC small abdominal aortic aneurysm (AAA) trial, was funded by the UK MRC and the British Heart Foundation with a series of grants. It was a multicentre RCT of the outcomes of early surgery compared with a period of ultrasound surveillance for patients aged 60–76 years with small asymptomatic AAAs of between 4 and 5.5 cm.

### Study details

Aortic aneurysm (a balloon-like swelling of the aorta) is a common condition; the prevalence in the population over 50 years is estimated from autopsy studies to be 3% and it is the third most common cause of sudden death. Nearly 1000 people die owing to aortic aneurysms in the UK each year, and a further 2000–3000 undergo successful aortic grafting each year. Mortality associated with elective surgery is now likely to be less than 5%, but post-emergency surgery mortality is high (40%) and numbers admitted for ruptured aortic aneurysms continue to rise.

Most surgeons would operate on an aneurysm over 6 cm in diameter, and few would operate on an aneurysm under 4 cm in diameter, owing to the different probabilities of rupture in each case. However, a grey area existed concerning asymptomatic aneurysms of sizes between 4 and 6 cm, and no consensus existed on the management of such AAAs.

The study used ultrasound scanning to determine the size of the asymptomatic infrarenal aneurysm. If it was less than 4 cm, the patient was observed, with a view to randomisation should the aneurysm grow. If it was between 4 and 5.5 cm and the patient was fit for surgery, then the patient was randomised into either an observation or a surgery group. If it was greater than 5.5 cm the patient was offered surgery. The observation group was seen every 6 months, but those with an aneurysm diameter between 5 and 5.5 cm were observed every 3 months.

An economic evaluation was undertaken. The costs of aortic grafting are considerable, as are the costs of regular observation. Furthermore, an unknown proportion in the observation group would require surgery for aneurysms that increased in size. Quality of life was also considered to be important, as those in the surgical group might be expected to return to normal quality of life within 6 months, while those in the observation group might display anxiety about possible aortic rupture, and as such might limit their activities. Quality of life was measured by the generic Medical Outcomes Study (MOS) instrument administered at the local level.

### Aims of the AAA trial

- To compare the mortality associated with early operation or observation in the management of small asymptomatic AAAs.
- To compare the costs and usage of NHS resources for patients having each treatment.
- To determine the growth and rupture rates of small AAAs.
- To investigate factors associated with the accelerated growth and rupture of small aortic aneurysms.
- To investigate, across Britain, how 24 hour and 3 day mortality rates for elective aneurysm surgery vary according to age, gender, aneurysm size and cardiovascular risk factors.

Each operation cost about £10,000 (in 1990), and so there exists an urgent need for rational decision-making with regard to the management of small asymptomatic AAAs. The results of the trial will dictate the future of aneurysms screening programmes and the management of small aneurysms.

## **Timing of the trial**

The funding application was submitted in August 1990 (economics funding application in February 1991) for a start date of January 1991 (April 1991) and expected completion in October 1998; it was published earlier than expected, in November 1998. First interviews were carried out in December 1998 and second interviews in July 2000.

## **Data from experts**

Four experts were approached for their opinions: two cardiovascular surgeons, a director of public health and a health economist with experience in undertaking economic evaluations of cardiovascular services. All agreed to participate, except for the director of public health who passed it to a junior colleague. The latter was thought inappropriate for the purposes of the study and, given the experience from the previous cases that a close familiarity with the technical and service aspects was necessary to address the questions precisely, and that this was a highly specialist topic, no further attempts were made to include a purchaser in the expert 'panel' for this study.

Between approaching the experts and carrying out the interviews the results were published in the *Lancet*. This happened earlier than the authors had been led to expect. A decision was made to proceed as before, sending them a set of four key papers<sup>64–67</sup> available at the time of funding and an outline of the trial, and asking the experts to answer the initial set of questions as at the time of funding, as for the previous cases, ignorant of the results of the trial (Appendix 6). The authors readily acknowledge the limitations of this suspension of knowledge. One month later the experts were sent the results<sup>68,69</sup> and asked the second set of questions (Appendix 7).

They were asked how, at the time of funding in 1990, they would have expected the treatment of small AAAs in patients aged 60–76 years to have developed had the trial not been funded, what

proportion would have undergone early elective surgery in 1990 and how would this have developed over the subsequent 10 years had the trial not proceeded.

All agreed that in 1990 almost all such aneurysms (some 95%) would have been treated by elective early surgery. With the development of ultrasound and the known mortality risk of elective surgery there would have been some slight movement towards more conservative management (watch and wait) over the next 10 years to possibly 10%, reducing those having surgery to about 90%. There might have been variation in surgery depending on age and size of aneurysm, with younger patients and those with larger aneurysms more likely to have received surgery. They thought that on the whole surgeons would think it safer to operate.

If the trial found in favour of early elective surgery this would confirm current and established practice, and all of the experts anticipated that this would lead to a further increase in the proportion receiving surgery, possibly up to 95%. The actual change would depend on the margin of difference between the procedures, and other developments such as progress in and costs of the new endovascular techniques. They thought that surveillance for smaller aneurysms would be likely to continue.

Were the conclusions of the trial 'inconclusive', all three thought that there would probably be little change in practice, unless there was an indication of no benefits of surgery. Response would depend on why the trial was inconclusive, and such a result might lead to more research, as RCTs are very unusual in surgery and there has been a high level of 'ownership' of this trial among clinicians. It was thought that this result would possibly lead to a move towards 15% surveillance and 85% surgery over the following 5 years.

Were the trial results to favour surveillance, all of the experts expected a change in practice towards surveillance. The impact would be less than if the result favoured surgery, as surgeons are more confident operating and it would be against the usual accepted practice. The shift was expected to be to 50% surveillance over the following 5 years.

### **Outcome measures**

The main outcome measure, which was all-cause mortality, was thought to be the right one and the most important outcome, although one consultant thought it unlikely to show a significant change as

deaths from AAA make up only 2.1% of all deaths and most of these are sudden death from undiagnosed aneurysms. All thought that quality of life was important, but did not think that the MOS was likely to be sensitive enough to pick up differences; it was thought that anxiety was the most likely to be significantly different between the groups, but was only a small component of MOS. However, it was thought that surgeons were not likely to be swayed by quality of life differences. All thought that health-service costs were important and that the health-service perspective was the right one to influence policy (patient costs were not taken into consideration).

#### **Threshold levels**

The experts were asked what barriers and threshold levels there might be concerning adoption of the more cost-effective management protocol, including what reduction in mortality, cost or quality of life might be necessary to trigger a change in practice. Two of the experts thought that a 5% difference and one thought that a 7–9% difference in mortality would trigger change. All thought that compared with mortality, costs would be considered of little importance unless they were over £2000 per patient. Quality of life differences would need to be very high at 20–30% differences.

#### **Probability of trial outcomes**

The experts were asked what they thought were the probabilities of the different outcomes of the trial. There was a considerable difference of opinion. The probability of an outcome in favour of surveillance was considered to be about 25%

(20%, 25% and 30%); in favour of surgery 5%, 20% and 60%; and inconclusive 20%, 50% and 70%. These ranges suggest that the experts were not influenced by or cognisant of the actual results; these ranges, together with equal weightings, will be used in weighting the results.

## **Model parameters, results and discussion**

It was decided that, with the rapid development of other techniques such as endovascular surgery, a period of 5 years would be appropriate for implementation. Therefore, the costs and benefits of treatment of patients presenting over the 5 years following the results, and their treatment and its outcome for a 5 year period (making 10 year follow-up altogether), were calculated and discounted to 1990. The cost per operation was taken from the literature as £10,000, and costs of outpatient ultrasound surveillance as £75 per visit. As the average rate of growth of an aneurysm of this size is about 0.5 cm per year and elective surgery would be offered for aneurysms reaching over 5.5 cm, it was estimated that half of the observation group would receive surgery after about 2.5 years. The size of the UK population aged 60–76 years with detected aneurysms of this size (in the absence of screening) is about 2000 per year. It is assumed that without the trial 90% would have surgery and 10% surveillance, with half of the latter requiring surgery within 5 years. *Table 14* sets out the scenario assumptions based on the experts' opinion.

**TABLE 14** Scenario assumptions

Trial outcome	Implementation scenario
<b>A</b> (favouring early surgery)	
Trial shows that early surgery is effective and cost-effective:	
<ul style="list-style-type: none"> <li>• mortality 29% surgery 36% surveillance</li> <li>• favourable quality of life</li> </ul>	<b>A</b> Increase in early surgery from 90% to 95% of presenting patients within 5 years: <ul style="list-style-type: none"> <li>• half of the surveillance patients would receive surgery within 5 years</li> </ul>
<b>B</b> (inconclusive)	
No significant difference in mortality or quality of life	
	<b>B</b> Following the trial, there is a small switch against surgery from 90% to 85% over 5 years: <ul style="list-style-type: none"> <li>• half of the surveillance patients would have surgery within 5 years</li> </ul>
<b>C</b> (favouring surveillance)	
No significant difference in mortality (29% both forms of care); 20% higher quality of life for surveillance patients:	
<ul style="list-style-type: none"> <li>• significantly lower costs for surveillance</li> </ul>	<b>C</b> Following the trial there is a shift against surgery from 90% to 50% over 5 years: <ul style="list-style-type: none"> <li>• half of the surveillance patients would have surgery within 5 years</li> </ul>

**TABLE 15** Cost and benefit calculations for the UK

	<b>Scenario A</b>	<b>Scenario B</b>	<b>Scenario C</b>
Trial result:			
Without trial			
Net cost (£ m)	45.457	45.457	45.457
Survivors	5442	5496	5496
Trial			
Trial cost (£ m)	0.987	0.987	0.987
Following trial (10 years)			
Cost (£ m)	46.616	44.699	41.110
Survivors	5454	5496	5496
Net trial implications			
Net costs (£ m)	2.146	0.229	-3.36
Lives saved	12	0	0
Cost per life saved	£178,800	-	-

**TABLE 16** Expected weighted costs and benefits (probability weightings)

	<b>Combination 1: Surgery 28% Survival 42%</b>	<b>Combination 2: Surgery 60% Survival 15%</b>	<b>Combination 3: Surgery 5% Survival 70%</b>	<b>Combination 4: Surgery 33% Survival 33%</b>
Expected costs (£ m)	-0.742	0.841	-2.187	-0.328
Expected lives saved	3.36	7.20	0.60	4.00
Expected cost/life	Dominant	£117,000	Dominant	Dominant

The costs of the trial and the costs and benefits of the implementation following each trial outcome were calculated using the assumptions above, and are presented in *Table 15* for these three scenarios. In each case the costs and benefits that would have occurred without the trial were identified in order to calculate the net implications of the trial, and the costs and benefits with and without the trial are presented for each scenario, along with the net costs and net benefits of the trial and finally the incremental cost per life saved.

An outcome favouring early surgery would result in an estimated net discounted cost of £2.1 million over 10 years for a net saving of about 12 lives, giving a cost per life saved of about £180,000. An inconclusive result is likely to lead to a net cost of £0.23 million over 10 years for no benefit of lives saved. An outcome favouring ultrasound surveillance would result in a net saving of some £3.36 million for some 20% quality of life benefit, but no significant net change in survival.

The net outcomes have been combined in a probabilistic framework to summarise the overall expected value of the trial. Probability weightings are based on the expert means, expert highest

values for surgery, expert lowest values for surgery and equal weightings for each scenario; weighted results are given in *Table 16*.

For combinations 1, 3 and 4 there is a small saving in costs of one-third to £0.85 million over 10 years, together with a saving of up to four lives. For combination 2, 60% probability favouring surgery, there would be a cost of £840,000 and an expected saving of some seven lives, giving an expected cost per life saved of about £117,000 (per life year of possibly £7500). The number of lives saved is not significantly different from zero in any of the scenarios.

## Conclusions

This case study report could represent advice to a research funding body about the likely returns to the proposed research. The conclusions relate to the likely payback in terms of costs and benefits, but also to the research design and issues around dissemination and likely implementation.

The main conclusion from this study is that the expected payback to the trial is marginally

positive, with expected small savings in both life and costs. Should the outcome be inconclusive there would be a likely cost of about a quarter of a million pounds for no benefit; a finding in favour of surveillance would result in a net discounted cost saving of about £3.4 million, while a finding in favour of early surgery would cost £2.1 million over 10 years, saving some 12 lives at a cost of £175,000 each, with an expected normal lifespan and quality of life, meaning a cost of about £20,000 per life-year saved.

Overall, the conclusion would be marginally in favour of the trial taking place, with a small but non-significant benefit from the trial. The benefits are likely to be low, partly because of the relatively small relevant population, the length of the trial and its cost, and because the cheaper care (surveillance) is unlikely to lead to a significant reduction in mortality but is likely to lead to surgery within 5 years for half of the patients. The length of the trial, although necessary, also reduces its likely relevance, as it may be overtaken by new surgical techniques of higher efficiency and lower cost. The design and outcome measures are considered appropriate, but a stronger measure of anxiety could be included.

## Actual results of the trial

The results of the trial were published in two papers in the *Lancet* in November 1998; one reporting the mortality results<sup>68</sup> and the other the health-service costs and quality of life.<sup>69</sup>

The main results were as follows:

- In total, 1090 patients aged 60–76 years with symptomless AAA 4.0–5.5 cm in diameter were randomised, 563 to undergo elective surgery and 527 ultrasonographic surveillance. Patients were followed for a mean of 4.6 years. Surgical repair was recommended if the diameter exceeded 5.5 cm in the surveillance group. The primary end-point was death. Analysis was by intention to treat.
- Both groups had similar cardiovascular risk at baseline; 93% of patients adhered to treatment and 309 patients died during follow-up.
- Mortality did not differ significantly between the groups at 2, 4 or 6 years. Age, gender and initial aneurysm size did not modify the overall hazard ratio.
- The overall hazard ratio for all-cause mortality in the early surgery groups compared with the surveillance group was 0.94 (95% CI 0.75 to

1.17,  $p = 0.56$ ). The 30 day operative mortality in the early surgery group was 5.8%, which led to a survival disadvantage for these patients early in the trial, countered by lower mortality later, so the cross-over point was at about 3 years.

- For patients in the surveillance group the median time to surgery was 2.9 years (when the aneurysm diameter had grown to more than 5.5 cm, increased by more than 1 cm per year, or was tender or ruptured).
- There was some indication of a non-significant higher mortality rate for those under 72 years in the surveillance group compared with the surgery group, but a lower mortality rate for those over 72 years. There was also a non-significant trend in relative mortality with diameter of aneurysm, favouring surveillance for aneurysms under 4.5 cm and surgery for those over 4.5 cm.
- The mean cost of treatment in the early surgery group was significantly higher than that for ultrasonographic surveillance, at £4978 versus £3914, a difference of £1064 (95% CI £799 to £1328). This finding was robust for a range of assumptions.
- Health-related quality of life was generally similar 12 months after randomisation for the two groups, but patients who underwent early surgery reported positive improvement in current health perceptions and less negative change in bodily pain.
- The research team conclude that ultrasonographic surveillance for small AAAs is safe, that early surgery does not provide a long-term survival advantage and that their results do not support a policy of open surgical repair for AAAs of 4.0–5.5 cm.

## Testing of the model

Copies of the two papers in the *Lancet* were sent to each of the experts, together with a second set of questions (Appendix 7). Of the three scenarios considered, the published outcome agrees closely with that favouring surveillance.

### **Ex post interviews with experts: their predictions of the effect of the trial results on subsequent policy implementation**

All three experts agreed with the authors that there should be a shift towards the use of ultrasound. One said that surveillance should be used for 5 years and then assessed nationally for mortality, rupture and natural history. He said that

the results may be overtaken by changes in technology to endovascular surgery down to 4 cm if that technique becomes cheaper. Another said that although some surgeons are saying this, it is illogical and not supported by the findings. One expert thought that the researchers' *Lancet* summary was somewhat misleading and that if surgeons wished to ignore their conclusions there is evidence in the paper that could be used to continue existing practice; that is, to use early surgery for younger patients and larger aneurysms, with surveillance for older patients and smaller aneurysms, and the extra cost of only £1000 would not be a likely restraint. However, all thought that the study would be highly respected and widely owned, as many of the UK vascular surgeons took part in it.

### Modelling *ex post* predicted outcomes

The *ex post* expert consensus was for a move towards more use of surveillance. However, it was not expected that this would be 100%, but more likely to be about 50% (from 10%) over about 5 years. This was consistent with their *ex ante* predictions for an outcome favouring surveillance.

If the move to surveillance was 50% over 5 years as the experts predicted, then the discounted 10 year cost saving would be £0.954 million, resulting in a non-significant net cost of some £0.033 million (−£0.271 million to £0.203 million). A move to a full 100% surveillance over 5 years would result in a discounted 10 year cost saving of about £2.416 million which, net of the cost of the trial, would result in a net saving of £1.43 million (£0.38 million to £2.03 million).

The *ex post* returns are less than predicted for the outcome, because the cost of surgical repair estimated during the trial was less than that estimated in the literature before funding. Overall, payback, estimated both *ex ante* and *ex post*, is low and non-significant owing to the relatively small population and the high proportion of those having surveillance later needing surgery.

In conclusion, the *ex post* predicted outcomes are very similar to the *ex ante* predictions and imply that the predicted 10 year savings will cover the cost of the trial by a small margin, but result in no significant change in expected mortality. It is unlikely that this study would have qualified for competitive funding on the grounds of likely payback.

## Chapter 7

# An application of the PATHS model: an assessment of the likely payback from a proposed UK trial of β-interferon for the treatment of multiple sclerosis

### Background

During the course of the study the authors received a request to apply their model to an application for funding for a large, randomised, double-blind multicentre trial, the UK-MS study: a trial of β-interferon (IFN-β) for the treatment of multiple sclerosis (MS), submitted to the HTA programme in the UK.<sup>70</sup> The results of this analysis are reported in addition to the case studies, as it provides an illustration of the use of the model.

The primary aims of the proposed trial were to evaluate the safety, efficacy and tolerability of IFN-β and of some other new treatments for MS. In addition, some limited information on resource use was to be collected 'to inform health economic analyses'. The trial was planned to last for 10 years. The UK HTA programme approached the authors to undertake an analysis of the likely payback from the proposed trial. This chapter outlines that assessment of the likely value of the proposed £20 million trial as carried out in January 1999.<sup>71</sup> It goes on to consider whether any subsequent evidence from other studies or changes in the policy-making environment might have changed the recommendations.

### Evidence available at the time

At the end of 1998, available evidence consistently pointed to some small clinical benefit to MS from IFN-β. Published trials had each added to the evidence that there does appear to be a statistically significant difference between standard treatment and IFN-β in terms of delayed progression and lower relapse rates.<sup>72–75</sup> The initial evidence was for relapsing remitting (RR-MS) patients, and later evidence suggested that there might also be benefit for secondary progressive (SP-MS) patients.<sup>75</sup> However, the clinical importance of these differences was uncertain (not least because of the difficulties in conducting therapeutic trials in this area),<sup>76</sup> as was

the effect of the observed differences on the quality of life of an MS patient compared with the effect of other, non-clinical interventions.

There had been no economic evaluation undertaken as part of a large randomised trial of IFN-β in the treatment of MS at the time the proposal was made. The debate on cost-effectiveness relied on modelling studies.<sup>77–82</sup> These studies had produced widely differing absolute values for the possible cost–utility ratio. Such variation is always likely to occur in cost-effectiveness ratios where the numerator (incremental cost) is large and the denominator (incremental effect) is very small. Despite this absolute variation, the various estimates of cost–utility ratios were consistent in policy terms in that none had, even in a sensitivity analysis, produced figures that would suggest that IFN-β for MS could be deemed cost-effective by normal health-service standards. To develop the evidence base further, we used the effectiveness data from the European Study on the use of IFN-β in SP-MS patients<sup>75</sup> in a published decision analysis model<sup>81</sup> to estimate the QALYs that would be gained from the costs of the trial and the likely levels of future prescribing that were judged would follow from the trial results.

A large definitive trial was proposed, although at that stage not precisely specified, to establish cost-effectiveness. The PATHS team was asked to evaluate the likely payback to such a trial. Summary results of this analysis and its conclusions are presented here.

### 'Payback' analysis and recommendation

Before examining the value of conducting a trial, it was necessary to examine the counterfactual: what was likely to happen to the level of prescribing during the next 10 years if the trial did not take place. The situation in the UK at that time was that only 1.5% of MS patients were

**TABLE 17** Principal possible outcomes of the trial of IFN- $\beta$ 

Trial outcome	Prescribing outcome
<b>A</b> No net overall clinical benefit and not cost-effective: • net treatment costs over 10 years: £78,000 • QALYs gained over 10 years: 0 (infinite cost per QALY)	<b>A</b> Following the trial, prescribing reduced to very low levels: • RR-MS: 280 patients • SP-MS: 400 patients
<b>B</b> Net overall clinical benefit but not cost-effective: • net treatment costs over 10 years: £78,000 • QALYs gained over 10 years: 0.07 (approx. £1 million per QALY)	<b>B1</b> Department of Health and health authorities impose stricter restraints on the prescribing of IFN- $\beta$ : prescribing is reduced in the light of cost-effectiveness information: • RR-MS: 1800 patients • SP-MS: 2500 patients <b>B2</b> Department of Health and health authorities impose no or weak restrictions on the prescribing of IFN- $\beta$ : IFN- $\beta$ is prescribed to a high proportion of those qualifying within the licence. • RR-MS: 12,000 patients • SP-MS: 10,000 patients
<b>C</b> Net overall clinical benefit and cost-effective; based on most optimistic assumptions: • net treatment costs over 10 years: £50,000 • QALYs gained over 10 years: 0.12 (approx. £400,000 per QALY)	<b>C</b> Prescribed to a higher proportion of those qualifying within the licence: • RR-MS: 16,000 patients • SP-MS: 14,000 patients

**Likelihood of outcomes**  
Outcome B: considered to be most likely (80% probability); B1 or B2 depends on Department of Health/NHS policy  
Outcomes A and C: considered to be unlikely (probability of 10% each)

receiving the drug,<sup>71</sup> in spite of well-organised information supplied by patient interest groups and of high patient expectation of the use of IFN- $\beta$ . This contrasted starkly with the levels of prescribing in most other European countries (6–23%) and the USA (16%). Prescribing was relatively very low in the UK owing to the terms of the UK licence, the need for prescribing by a consultant neurologist and local allocation of very limited funds for IFN- $\beta$ .

Expert clinicians, an economist and a statistician/epidemiologist, all of whom had worked closely in this research area, were consulted to inform the assumptions.

It was hypothesised that future levels of prescribing in the UK might be altered by:

- a widening of the terms of the UK licence following the results of, what were at that time, recently published and ongoing trials showing

benefit for other subgroups of MS patients in addition to RR-MS patients

- the expiration of the current patent for IFN- $\beta$ , and the possibility of lower price generic products
- the availability of new products for treating MS, some of which might be used in combination with, or as an alternative to, IFN- $\beta$
- most importantly, the policy position of the Department of Health and of individual commissioners of healthcare that might be influenced by any or all of the three factors listed above.

It was concluded that the above pressures were likely to lead to an upward trend in prescribing rates. However, the size of this increase would ultimately depend on the continuing determination of central and local policy makers to act to constrain the availability of IFN- $\beta$  in the absence of firm evidence of an acceptable cost-effectiveness ratio.

**TABLE 18** Cost and QALY calculations for principal possible outcomes

	<b>Scenario A</b>	<b>Scenario B1</b>	<b>Scenario B2</b>	<b>Scenario C</b>
Without trial (20 years)				
Net therapy cost (£ m)	725	725	725	465
Benefits	0	651	651	1115
During trial (10 years)				
Trial cost (£ m)	20	20	20	20
Net therapy cost (£ m)	273	273	273	175
Benefits (QALYs)	0	245	245	420
Following trial (10 years)				
Net therapy cost (£ m)	53	335	1716	1500
Benefits (QALYs)	0	301	1540	3600
Net trial implications				
Net costs (£ m)	-379	-97	1284	1230
Net benefits (QALYs)	0	-105	1134	2095
Cost/QALY of trial and resulting prescribing	No change in QALYs, but saving of £379 m	£924,000	£1.13 m	£423,000
Judgement of probability of outcome	10%	80% B1 or B2 depends on Department of Health policy		10%

The potential payback from the proposed trial was considered against this background using previously published frameworks of Buxton and Hanney<sup>9</sup> and Townsend and Buxton.<sup>16</sup> The latter, a precursor to the PATHS model as described in this report, used a quantitative analysis to assess the likely value of a proposed MRC trial of long-term HRT. It examined three main scenarios, reflecting different possible trial outcomes and their resulting prescribing levels. Building on this and the early development of PATHS, we similarly looked at three possible outcomes for the MS trial and what would be likely to follow from them in terms of clinical policy for MS in the UK over a 20 year period (10 year trial and 10 years post-trial). (Full details of the calculations summarised here are contained within the original report to the UK HTA programme.<sup>71</sup>)

It was assumed, based on the expert opinion, that if the trial were not to proceed, then the number of MS patients prescribed IFN-β would increase over time. The model assumed that over 10 years, the number of RR-MS patients receiving IFN-β would increase from the then current position (approximately 1200 patients) to 2400 patients, and that this level would remain constant for the subsequent 10 years. It was also assumed that IFN-β would be gradually made available to 3400 patients with SP-MS over the next 10 years, and again that this level would hold over the subsequent 10 years.

The principal possible outcome scenarios of the proposed trial are set out in *Table 17*, and the costs and benefits for each with or without the trial (the counterfactual) in *Table 18*.

For each of these scenarios estimates were made of the costs of the trial, the benefits that would arise during the trial, the cost of the therapy following the trial and the health benefits that would arise from that. In each case the costs and benefits that would have arisen without the trial were netted off. This gives the net implications of the trial, which are presented in terms of net costs, net QALYs and the implied cost per QALY of the trial and its resulting prescribing (*Table 18*).

Under only one of these scenarios can the proposed trial be justified on cost-effectiveness grounds. This (scenario A) makes the unlikely assumption that the trial shows IFN-β to have no net benefit to patients, and that as a result its prescribing would be discontinued.

In scenario A the trial has no effect on QALYs (the intervention is found to be ineffective), but leads to a net saving from a subsequent reduction in prescribing. Scenario B1 results in a saving from undertaking the trial of just under £100 million, which is offset by a reduction in benefits to patients of 105 QALYs (equivalent to a cost saving per QALY foregone of approximately £900,000). In contrast, scenario B2 (with a weak policy)

results in an increase in costs of £1284 million, to be offset against 1134 QALY gains following the trial. Therefore, the cost per QALY of the trial in this situation is £1.13 million. Finally, if scenario C were to arise, in which the cost-effectiveness of the intervention were shown to be much more favourable, then the cost per QALY of the trial and subsequent prescribing would still be over £400,000.

Arguably, the least useful outcome of the trial would be scenario B: the production of strong independent evidence of some patient benefit, in combination with further, stronger evidence that IFN-β is not cost-effective. The report suggested that this outcome might leave policy makers in the difficult situation of being unwilling to act on the latter because of the former. Based on expert opinion, and in line with all current evidence,<sup>82–84</sup> the most likely outcome of the proposed trial was judged to be B: that the trial would show some net overall clinical benefit at a very considerable net cost.

Each of these scenarios was an attempt to consider what might happen from a broad range of detailed possibilities. However, the analysis indicated that the policy conclusions were not sensitive to detailed assumptions. The only attractive (cost-effective) outcome was one in which the evidence was such that it could lead to a reduction in prescribing. This could follow from evidence that there was no net overall clinical benefit or that the net effect for patients was negative. Neither seemed likely given existing evidence. The question then became whether a trial producing stronger evidence of a small effect at a very high cost per QALY would lead to a reduction or an increase in prescribing. This would depend on the policy response from the government department.

The benefits of the proposed trial would be highly dependent on the policy adopted both during and after the trial, and the extent to which there was willingness to restrain prescribing of IFN-β based on cost-effectiveness information. Unless prescribing restraint could be firmly maintained the trial would probably not be feasible.

None of the levels of cost per QALY that were calculated approached conventionally acceptable levels. Indeed, on the available evidence no likely way could be seen in which this trial and its resulting impact on therapy could achieve a conventional threshold level of cost per QALY. The most favourable net cost per QALY of the trial plus resulting prescribing exceeded £400,000

per QALY, and was based on the scenario that the trial would find strong evidence of a net overall clinical benefit and that IFN-β would be found to be much more cost-effective than suggested by any existing analysis. The likelihood of this outcome was judged to be no more than 10%. The more likely scenario suggested a cost per QALY of around £1 million.

If the current evidence base was inadequate for policy, it was important to consider what were the main areas of uncertainty that needed to be addressed. It was clear that a trial to increase the precision of the estimate of small differences in clinical effectiveness was not needed. Rather, any trial to improve the evidence base for policy makers needed to address the key policy question of cost-effectiveness. There was no need for precision about how effective IFN-β was, if under all plausible circumstances, it was not likely to be cost-effective.

The assessment also considered other categories of possible payback.<sup>9</sup> In particular, it was recognised that the proposed trial would generate knowledge benefits of wider relevance to the understanding of MS, particularly on disability and quality of life, of the non-healthcare costs associated with MS, and information to assist in the interpretation of the long-term implications of short-term indicators of disease progression.<sup>82</sup> Although each of these would be of value in interpreting future research, it was argued that each could be achieved through other less expensive studies.

## **Subsequent research and policy decisions**

The UK HTA programme did not fund the proposed trial. It did, however, commission, in collaboration with other interested parties including the MS Society and commissioners of healthcare, five reviews, the first four of which have been published, of the evidence on:

- disease-modifying drugs<sup>83</sup>
- the natural history and epidemiology of MS: modelling of the burden of morbidity and disability<sup>85</sup>
- the role of specialist nurses<sup>84</sup>
- treatment for fatigue<sup>86</sup>
- the management of pain and spasticity (not yet published).

Each of these reviews is being considered by the National Institute for Clinical Excellence (NICE)

as it prepares its guidelines for the management of MS.

In the meantime, NICE was asked in August 1999 by the Department of Health and the National Assembly for Wales to appraise IFN- $\beta$ /glatiramer for MS. There followed a prolonged process of appraisal, appeal and reappraisal, which included the commissioning of further economic modelling in the context of the Appraisal Committee's doubts about the cost-effectiveness of these drugs. Finally, early in 2002, NICE issued its appraisal determination, which confirmed their view that the drugs could not be recommended because of their poor cost-effectiveness.<sup>87</sup> Based on the additional modelling, NICE reported the cost per QALY gained over a 20 year time frame to be between £40,000 and £90,000. The Department of Health responded to this conclusion by announcing a 'risk-sharing' arrangement that would, following reductions in the price of the drugs, make the drug available to appropriate patients, while those patients would be monitored to 'confirm' the cost-effectiveness of the drugs.<sup>88</sup> If this cost-effectiveness did not reach a pre-agreed level, then the companies would have to reduce further the costs to the NHS.

## Conclusions

The conclusion was that it was highly unlikely that IFN- $\beta$  for MS was cost-effective, although it was

conceivable that the drug could be cost-effective if precisely targeted at groups of patients who would be most responsive to the drug, if such could be identified.<sup>89-91</sup> Forbes and co-workers<sup>92</sup> looked at cost utility in a possible subgroup of ambulatory SP-MS patients with more active disease, but still estimated a cost per QALY of over £800,000 (95% CI £161,000 to infinity). The subsequent detailed cost-effectiveness modelling has produced lower estimates of the cost per QALY but, at the then prevailing prices, still not within normally accepted limits for the NHS.

It was also concluded that the proposed trial would not be cost-effective, and the analysis reminds us that we should no more presume that a technology assessment will be cost-effective than will a technology itself. HTA funding should be directed to those areas in which timely results are likely to influence policy making. This requires an analysis that considers both the theoretical value of information and the likely behavioural response to it. Formal Bayesian analysis of the value of information may offer a way to improve our estimates of the former; better political, science-based understanding of what influences policy making may help to improve our estimates of the latter.



## Chapter 8

# Conclusions

The PATHS study set out to synthesise a methodology for assessing the likely payback or returns to proposed health technology assessments and to apply this methodology in case studies. It was recognised from the outset that a full formal testing of the methodology was impracticable in terms of timescale and resources. Such a formal test would require assessment before, but independent of and hidden from the decision to fund a series of proposed projects, and then follow-up of the actual payback from these projects over an extended period after the completion and dissemination of the research. To assess the impact of the method in practice would involve an even more extensive and elaborate study, comparing a research programme that used the proposed methodology with one that did not. This would take several decades. The more modest aim here is, therefore, to demonstrate the potential value of a practical and feasible approach, and to indicate the sort of impacts that it might be expected to have.

The approach developed here emphasises the benefits from the implementation of the results of assessment, and estimates the costs and benefits that may flow from alternative outcomes of a research project or an area of research. Expert opinion is used to inform the decision analysis model, in a way that builds on and formalises the usual process of incorporating expert input from potential users of research into the decision-making about the research funding. It uses this to consider the likelihood that results from a specific proposed study will influence policy, and the nature of the likely policy change.

For two of the three case studies undertaken, the model predicted the likely payback well, in that the *ex ante* and *ex post* analyses both indicated either clinical benefit or no clinical loss of benefit, together with health services cost savings. The result for the third case was inconclusive owing to a very high (45%) level of non-compliance with protocols, which meant that the potential benefits of that research project were not realised and so did not provide a test of the model. In the live example given of an application of the model to inform the decision whether a proposed study should be carried out,<sup>71</sup> the result appeared robust

and informed the decision not to fund a proposed trial. It may be concluded that the model has a valuable part to play in the research prioritisation process, alongside existing criteria.

The strength of the PATHS approach is its emphasis on the effect of possible outcomes of the research on policy and/or practice, to an extent that other models do not. It also has the advantages of being transparent and open to sensitivity analysis. Refereeing or reviewing procedures for research proposals typically focus on scientific merit, as judged by peer reviewers, who are likely to be other academic researchers. This tends to focus attention correctly on the scientific criteria of validity, power and internal consistency, and encourages researchers to use as end-points measures or instruments that are well validated and for which there are existing data on which to base power calculations. However, where policy makers respond to research outcomes with a view to implementation, they may have very different concerns, and be far more concerned about the clinical or policy relevance and importance of the outcomes measured. 'Hard' end-points such as mortality are likely to have equal validity and relevance for researcher and policy maker or commissioner alike. Softer endpoints, such as dimensions of the SF-36, may have high validity and credibility with researchers, but may have little resonance with policy makers, compared with say cost or cost-effectiveness or patient satisfaction. *Ex ante* modelling informed by expert panels, as demonstrated here, not only can help to identify where there is likely to be substantial payback to research, but also may be useful in indicating how the research design might be improved to increase its chance of impacting on policy and practice and so increase cost-effectiveness. Equally, the process may help to identify the most effective means of disseminating knowledge of an ongoing trial and its results, to facilitate implementation and change of practice.

The involvement of experts is a strength but also a limitation to the model, and raises a number of questions that could not be addressed directly. No formal assessment was made of the optimal combination of experts to consult or the effect of

increasing their number. We believe that the key criterion is that they have sufficient detailed expertise to have and to retain a realistic view of what will change policy in the particular area of interest. There was a considerable degree of consensus between the experts, but in future research it would be important to assess how sensitive the results might be to reasonable differences in expert opinion.

A number of important issues for payback have been thrown up by the research. The UK NHS R&D programme, among others, has set relevance to the improvement of health and health services as the keystone for research prioritisation for which this model would be very appropriate. However, if this is to be assessed it is very important to know the baseline level of use of the technologies or services being evaluated or compared. This was not known or included in the research proposals for any of the test cases, although it was known for the live application. Surveys and audits tend not to be valued in research, although a case could be made that survey data are an essential adjunct to a literature review, for assessing the relevance and potential importance of a health technology assessment. In the NHS R&D system, such survey data, where known, have usually formed an element of the 'vignettes' produced as part of the prioritisation process. In these, such survey data are used to set out the size of the problem, but often the detail required is not available. Information on the current use of a technology, and its expected trajectory, is essential to the *ex ante* assessment of payback. Payback cannot be evaluated without an assessment of the counterfactual; that is, what would happen regarding the use of the technology if the research did not take place. Implications for payback are very different for a new, as yet unused technology that is likely to be adopted only if good evidence is provided, compared with a technology which, despite the lack of evidence, is in increasing or widespread use. A large part of the payback in the cases considered here was due to an expected reduction in the use of the technology if the research showed it to have low relative benefit. Negative results may produce a high payback. The process also requires explicit consideration of the length of time over which the research is likely to influence policy. This will reflect a number of factors, such as the timing and credibility of other studies, including international studies, and likely changes to the technology or its competitors. In an area of rapid technological change, the policy relevance of research may be quite transient.

The issue of equipoise is an interesting and ironic one. There is a strong emphasis in the literature on the need for equipoise as the ethical basis for an RCT. However, the finer the state of equipoise, the larger the study and the greater the risk of an inconclusive result. An inconclusive result is likely to give low returns to the research funding and not to be a good use of resources, although there is likely to be some gain in knowledge. A conclusive result should lead to either an increase in health benefit or a reduction in costs, and is likely to repay the research costs unless these are very high compared with the cost of using the technology. Where the returns are assessed as positive, the level of return will need to be considered against returns to other competing proposals in the prioritisation process. It is unlikely that sufficient funds would be available to fund all research with an acceptable incremental cost-effectiveness ratio.

The model developed herein drew on existing literature and approaches, particularly those proposed by Weinstein,<sup>25</sup> Eddy,<sup>22</sup> and Townsend and Buxton.<sup>16</sup> Given the elapsed time since the start of the project, it may be helpful to recontextualise this work in the light of other research published since the original literature review. The literature search undertaken in May 1998 was repeated in January 2002, and it identified 276 potential articles published in the intervening period. Of these, only 12 appeared possibly of direct relevance. Based on the full text of these, three papers were found that had made important contributions which, had they been published earlier, would have figured significantly in our consideration of the best methodology to adopt.

The first study is the work of Davies and colleagues,<sup>93</sup> which attempted a similar task to that reported here. Davies and colleagues attempted to apply a decision-analytical model, based on their earlier work<sup>21</sup> that was reviewed in Chapter 2, to several topics considered by the UK HTA programme in 1997 and 1998. Their approach is in many ways similar to our own, except that they focus entirely on the consideration of a topic area, and do not consider specific research proposals to study that topic. Their approach attempts to provide a formal overall prioritisation at an early stage in the research commissioning sequence. They do not attempt to address the question of the extent to which a particular study will or will not be convincing to policy makers. Their approach does, however, recognise that policy and practice do not

automatically adjust to reflect the changing evidence, and it makes assumptions about the level of likely uptake of the technology with and without evidence of its effectiveness. It implicitly assumes that all research has clear results, which is not necessarily so, although it potentially allows for the possibility of false-positive or false-negative results. Davies and colleagues offer a guarded conclusion that such an approach is feasible in specific contexts, but needs further testing and validation of results.

Claxton and others have also taken forward earlier proposals<sup>17,32,33</sup> reviewed in Chapter 2, for undertaking formal Bayesian 'value of information' analyses to inform decision-makers about the expected value of reducing uncertainty about specific parameters in health technology assessment contexts. This further work includes an application that assesses the maximum value of additional research in the context of the treatment of Alzheimer's disease.<sup>94</sup> This work and other Bayesian applications (e.g. by Fryback and colleagues<sup>95</sup> and Briggs<sup>96</sup>) are beginning to demonstrate quite clearly the technical feasibility of such analyses, although formal Bayesian approaches remain opaque and inaccessible to most decision-makers (see Sheingold<sup>97</sup>). They can clearly provide important information and insights concerning the design of major clinical trials and economic evaluation studies, but are still currently very demanding in terms of the scarce analytical skills available to undertake them, and the sparse understanding of the techniques by potential users. Perhaps most importantly, they do not consider the likely policy response to additional information, but assume a totally rational model of behaviour.

The third paper, by Meltzer,<sup>98</sup> suggests alternatives to the EVPI, including maximum value of research, which are less demanding in terms of knowledge of priors, and may offer a first stage approach that deserves further consideration.

Finally, to address the issue of the uncertain impact of research on practice, Lilford and colleagues,<sup>6</sup> have suggested an approach to research commissioning that involves potential research users directly in the formulation of the details of the research. They propose a method of research in which the scope, form and content of research on the delivery and organisation of health services are developed iteratively by a research director advised by a commissioning group of health-service managers and research commissioners. The aim is to provide flexibility in

responding to changing circumstances and, by engaging potential users of the research in its production, to enhance the likelihood of their putting it to use. They emphasise that such evidence as exists suggests that scientists need to encourage decision-makers and be engaged by them if the results of research are to impact on practice. This form of involvement of users is focused less on prioritisation of the 'right' topic and more on ensuring that the research is developed in ways that are most likely usefully to inform policy and practice and which might, taking the example used earlier, include the relevance of alternative possible end-points with relevance for policy.

## **Recommendations to HTA funders**

The PATHS model demonstrated here offers a conceptually simple approach to prioritising research in terms of its overall likely net effect on health benefits and costs of implementation, as well as the costs of the assessment itself. It considers the various likely outcomes of the research and their probabilities. It features the use of expert opinion from policy makers, clinicians and academics to increase the relevance of the assessment, and to supplement the knowledge of the likely alternative outcome scenarios on policy and implementation. It not only provides an assessment of the cost-effectiveness of the research, but may identify ways in which the research design, end-points, analytical methods or dissemination can be enhanced.

It is important that the type of analysis proposed here be carried out by competent and impartial evaluators and that it is transparent. The cost of the sort of analyses that were used here is likely to be between 1 and 4 weeks' work for a researcher, depending on the information available and the complexity of the proposal. There may also be costs of a small honorarium to the 'expert' clinician, manager or health economist for their time. A total cost in the range of £1000 to £4000 per project or area is likely, possibly higher for something large and complex. This would represent a small proportion of the proposed research cost and would be likely to give a good return in terms of excluding low return proposals or improving relevance and likelihood of implementation. This is particularly true where an outcome is likely to lead to a non-marginal change in an expensive technology or to go against the current trend. It may best be used after

preliminary prioritisation has taken place and as an important contribution to the final decision to fund.

The authors recommend that formal analysis of potential payback, along these lines, should be undertaken as part of an ongoing evaluation for projects costing over a certain threshold of, say, £250,000. For very expensive projects, some formal value of information analysis might also be routinely appropriate. The cost of the process demonstrated here, if undertaken routinely, would be small in comparison to the cost of funding the research. Whether a threshold were used or not, the scale and intensity of the exercise could be varied, to reflect the cost, perceived policy importance and contentiousness of the proposal.

## **Recommendations for further research**

Neither the PATHS model nor other current models provide a complete solution to the

problems of prioritising topics and research proposals. Together, they point towards possible practical strategies and to a further research agenda in which characteristics of all strategies might be incorporated. Therefore, we think that there is a high priority for research to synthesise these approaches.

This investigation was able to compare only immediate *ex post* assessment of likely implementation with the *ex ante* situation. Long-term follow-up of actual implementation is fundamental to formal testing of this approach and research is needed in this area.

Expert opinions have a key role to play in the model and research is required to assess how robust the approach is to the choice and number of experts.



## Acknowledgements

We should like to thank the Medical Research Council and the HTA programme for advice on possible funded research to include as cases, the researchers for permission to use their

research as cases, and the expert clinicians, economists, epidemiologists, academics, managers and purchasers who gave us their expert opinions.



## References

1. Donaldson MS, Sox HC, editors. Setting priorities for health technology assessment. A model process. Washington DC: National Academy Press; 1992.
2. Office of Science and Technology and Programme for Policy Research in Engineering Science and Technology. Returns to research and development spending. London: HMSO; 1993.
3. NHS Executive. National Co-ordinating Centre for Health Technology Assessment. The Annual Report of the NHS Technology Assessment Programme 1997. Identifying questions, finding answers. NHS Executive; 1997.
4. Henshall C, Oortwijn W, Stevens A, Granados A, Banta D. Priority setting for health technology assessment. Theoretical considerations and practical approaches. A paper produced for the EUR-ASSESS project. *Int J Technol Assess Health Care* 1997;13:144–85.
5. Harper G, Townsend J, Buxton M. The preliminary economic evaluation of health technologies for the prioritization of health technology assessments: a discussion. *Int J Technol Assess Health Care* 1998; 14:652–66.
6. Lilford R, Jeeock R, Shaw H, Chard J, Morrison B. Commissioning health services research: an iterative method. *Journal of Health Services Research and Policy* 1999;4:164–7.
7. Ham C, Hunter DJ, Robinson R. Evidence based policy making. *BMJ* 1995;310:71–2.
8. Pickles H. Research and development in the new NHS: another challenge for the profession. *J R Coll Phys Lond* 1996;30:509–11.
9. Buxton M, Hanney S. How can payback from health services research be assessed? *Journal of Health Services Research and Policy* 1996;1:35–43.
10. Buxton M, Hanney S. Evaluating the NHS research and development programme: will the programme give value for money? *J R Soc Med* 1998;91 (Suppl 35):2–6.
11. Black N. A National Strategy for Research and Development: lessons from England. *Annu Rev Public Health* 1997;18:485–505.
12. Lilford RJ, Braunholtz DA, Greenhalgh R, Edwards SJL. Trials and fast changing technologies: the case for tracker studies. *BMJ* 2000;320:43–6.
13. Hanney S, Davies A, Buxton M. Assessing the benefits from health research projects: can we use questionnaires instead of case studies? *Research Evaluation* 1999;8:189–99.
14. Buxton M, Hanney S, Packwood T, Roberts S, Youll P. Assessing benefits from Department of Health and National Health Service research and development. *Public Money and Management* 2000;20:29–34.
15. Croxson B, Hanney S, Buxton M. Routine monitoring of performance: what makes health R&D different? *Journal of Health Services Research and Policy* 2001;6:226–32.
16. Townsend J, Buxton M. Cost effectiveness scenario analysis for a proposed trial of hormone replacement therapy. *Health Policy* 1997;39:181–94.
17. Claxton K, Posnett J. An economic approach to clinical trial design and research priority-setting. *Health Econ* 1996;5:513–24.
18. Detsky AS. Using economic analysis to determine the resource consequences of choices made in planning clinical trials. *J Chron Dis* 1985;38:753–65.
19. Detsky AS. Are clinical trials a cost-effective investment? *JAMA* 1989;262:1795–800.
20. Detsky AS. Using cost-effectiveness analysis to improve the efficiency of allocating funds to clinical trials. *Stat Med* 1990;9:173–84.
21. Drummond MF, Davies LM, Ferris FL III. Assessing the costs and benefits of medical research: the Diabetic Retinopathy Study. *Soc Sci Med* 1992;34:973–81.
22. Eddy DM. Selecting technologies for assessment. *Int J Technol Assess Health Care* 1989;5:485–501.
23. Phelps CE, Parente ST. Priority setting in medical technology and medical practice assessment [published erratum appears in *Med Care* 1992;30:744–5]. *Med Care* 1990;28:703–23.
24. Phelps CE, Mooney C. Correction and update on ‘priority setting in medical technology assessment’. *Med Care* 1992;30:744–51.
25. Weinstein MC. Cost-effective priorities for cancer prevention. *Science* 1983;221:17–23.
26. Institute of Medicine. Setting priorities for health technology assessment: a model process. Washington, DC: National Academic Press; 1990.
27. Shachtman RH. Decision analysis assessment of a national medical study. *Operations Research* 1980;28:44–59.

28. Thompson MS. Decision-analytic determination of study size. The case of electronic fetal monitoring. *Med Decis Making* 1981;1:165–79.
29. Torgeson D, Donaldson C, Reid D. Using economics to prioritize research: a case study of randomized trials for the prevention of hip fractures due to osteoporosis. *Journal of Health Services Research and Policy* 1996;1:141–46.
30. Wennberg J, Gittlesohn A. Small area variations in health care delivery. *Science* 1973;182:1102–8.
31. Sculpher MS, Drummond MF, Buxton M. The iterative use of economic evaluation as part of the process of health technology assessment. *Journal of Health Services Research and Policy* 1997;2:26–30.
32. Claxton K. The value of information in the economic evaluation of new health care technologies: an application to a policy model of Alzheimer's disease. *Med Decis Making* 1998;18:4.
33. Claxton K. Bayesian approaches to the value of information: implications for the regulation of new pharmaceuticals. *Health Econ* 1999;8:269–74.
34. Lilford R, Royston G. Decision analysis in the selection, design and application of clinical and health services research. *Journal of Health Services Research and Policy* 1998;3:159–66.
35. Sheldon TA. Problems of using modelling in the evaluation of health care. *Health Econ* 1996;5:1–11.
36. Office of Technology Assessment. Research funding as an investment: can we measure the returns? Washington, DC: Government Printing Office; 1986.
37. Smith K. Economic returns to R&D: methods, results and challenges. Review Paper No. 3. London: Science Policy Support Group; 1991.
38. Hutton J, Backhouse R. Payback from research and development: final report. London: Battelle Medical Technology and Policy Research Centre; 1994.
39. Oortwijn WJ, Vondeling H, Bouter L. The use of societal criteria in priority setting for health technology assessment in the Netherlands. Initial experiences and future challenges. *Int J Technol Assess Health Care* 1998;14:226–36.
40. Briggs A. Handling uncertainty in the results of economic evaluation. *Office of Health Economics Briefing* 1995;32:1–12.
41. Writing Group for the PEPI Trial. Effects of estrogen or estrogen/progestin regimens on heart disease factors in post menopausal women. *JAMA*, 1995;273:199–208.
42. Kiel DP, Felson DT, Anderson JJ, Wilson PW, Moskowitz MA. Hip fracture and the use of oestrogens in post menopausal women: the Framingham Study. *N Engl J Med* 1987;317:1169–74.
43. Griffin J. Osteoporosis. London: Office of Health Economics; 1990.
44. Colditz GA, Hankinson SE, Hunter DJ, Willet WC, Manson JE, Stamper MJ, et al. The use of estrogen and progestins and the risk of breast cancer in postmenopausal women. *N Engl J Med* 1995;322:1589–93.
45. Daly E, Roche M, Barlow D, Gray A, McPherson K, Vessey M. Hormone replacement therapy: an analysis of benefits, risks and costs. *Br Med Bull* 1992;48:368–400.
46. Daly E, Gray A, Barlow D, McPherson K, Roche M, Vessey M. Measuring the impact of menopausal symptoms on quality of life. *BMJ* 1993;307:836–40.
47. Townsend J. Hormone replacement therapy: assessment of present use, costs and trends. *Br J Gen Pract* 1998;48:955–8.
48. US Congress, Office of Technology Assessment. The menopause, hormone therapy and women's health (OTA-BP-BA-88). Washington, DC: US Government Printing Office; 1992.
49. Garcia J, Renfrew M, Marchant S. Postnatal home visiting by midwives. *Midwifery* 1994;10:40–3.
50. Glazener C, Abdalla M, Russell I, Templeton A. Postnatal care: a survey of patients' experiences. *British Journal of Midwifery* 1993;1:67–74.
51. Twaddle S, Liao XH, Fylie H. An evaluation of postnatal care individualised to the needs of the woman. *Midwifery* 1993;9:154–60.
52. Van Teijlingen ER. The profession of maternity home care assistant and its significance for the Dutch midwifery profession. *Int J Nurs Stud* 1990;27:355–66.
53. Morrell CJ, Spiby H, Stewart P, Walters S, Morgan A. Costs and benefits of community postnatal support workers: a randomised controlled trial. *Health Technol Assess* 2000;4(6).
54. Deakin CD, Hicks IR. AB or ABC: pre-hospital fluid management in major trauma. *Journal of Accident and Emergency* 1994;11:154–7.
55. Deakin C, Davies G. Defining trauma patient subpopulations for field stabilization. *Eur J Emerg Med* 1994;1:31–3.
56. Deakin CD. Early fluid resuscitation in haemorrhagic shock. *Eur J Emerg Med* 1994;1:83–5.
57. Aprahamian C, Thompson BM, Towne JB, Darin JC. The effect of a paramedic system on mortality of major open intra-abdominal vascular trauma. *J Trauma* 1983;23:687–90.
58. Pons PT, Moore EE, Cusick JM, Brunko M, Antuna B, Owens L. Pre-hospital venous access in an urban paramedic system – a prospective on-scene analysis. *J Trauma* 1988;28:1460–3.

59. Smith JP, Bodai BI, Hill AS, Frey CF. Pre-hospital stabilization of critically injured patients: a failed concept. *J Trauma* 1985;25:65–70.
60. O'Gorman M, Trabulsky P, Pilcher DB. Zero-time pre-hospital IV. *J Trauma* 1989;29:84–6.
61. Wears RL, Winton CN. Load and go versus stay and play: analysis of pre-hospital IV fluid therapy by computer simulation. *Eur J Emerg Med* 1994; 1:31–3.
62. Netten A, Curtis L. Unit costs of health and social care. Canterbury: Personal Social Services Research Unit, University of Kent; 1999.
63. Turner J, Nicholl J, Webber L, Cox H, Dixon S, Yates D. A randomised controlled trial of pre-hospital intravenous fluid replacement therapy in serious trauma. *Health Technol Assess* 2000;4(31).
64. Bengtsson H, Bergqvist D, Jendteg S, Lindgren B, Perrson U. Ultrasonographic screening for abdominal aortic aneurysm: analysis of surgical decisions for cost-effectiveness. *World J Surg* 1989;12:266–71.
65. Cheatle TR, Scurr JH. Abdominal aortic aneurysms: a review of current problems. *Br J Surg* 1989;76:826–9.
66. Collin J, Araujo L, Walton J, Lindsell D. Oxford screening programme for abdominal aortic aneurysm in men aged 65 to 74 years. *Lancet* 1998; ii:613–15.
67. Quill DS, Colgan MP, Summer DS. Ultrasonic screening for the detection of abdominal aortic aneurysms. *Surg Clin North Am* 1989 Aug;69:713–20.
68. The UK Small Aneurysm Trial Participants. Mortality results for randomised controlled trial of early elective surgery or ultrasonographic surveillance for small abdominal aortic aneurysms. *Lancet* 1998;352:1649–55.
69. The UK Small Aneurysm Trial Participants. Health service costs and quality of life for early elective surgery or ultrasonographic surveillance for small abdominal aortic aneurysms. *Lancet* 1998; 352:1656–60.
70. A proposal for a randomised controlled trial of beta interferon and other disease-modifying treatments for MS. The UK-MS study. UK HTA programme, Confidential submission Sept 1998.
71. Buxton M, Robert G, Townsend J, Harper G. An assessment of the likely 'payback' from the proposed UK-MS trial of beta interferon. Uxbridge: HERG, Brunel University; 1999.
72. IFNB Multiple Sclerosis Study Group. Interferon beta-1b is effective in relapsing-remitting multiple sclerosis. I. *Neurology* 1993;43:655–66.
73. Jacobs LD, Cookfair DL, Rudick RA, Herndon RM, Richert RJ, Salazar AM, et al. Intramuscular interferon beta-1a for disease progression in relapsing MS. The MS Collaborative Research Group. *Ann Neurol* 1996;39:285–94.
74. PRISMS Study Group. Randomised double-blind placebo-controlled study of interferon beta-1a in relapsing/remitting MS. *Lancet* 1998;352:1498–504.
75. European Study Group on Interferon β-1b in Secondary Progressive MS. Placebo-controlled multicentre randomised trial of interferon β-1b in treatment of secondary progressive MS. *Lancet* 1998;352:1491–7.
76. Goodin DS. Perils and pitfalls in the interpretation of clinical trials: a reflection on the recent experience in multiple sclerosis. *Neuroepidemiology* 1999; 18:53–63.
77. Nicholson T, Stein K. Interferon beta-1a in relapsing remitting MS. Development & Evaluation Committee Report No. 77; December 1997. Wessex Institute, Southampton.
78. Otten N. Interferon beta-1b and MS. CCOHTA Technology Overview: Pharmaceuticals, Issue 5.0; December 1996. CCOHTA, Ottawa.
79. Otten N. Comparison of drug treatments for multiple sclerosis. CCOHTA Technology Report; December 1998. CCOHTA, Ottawa.
80. Richards RG, McCabe CJ, Cooper NJ, Paisley SF, Brennan A, Akehurst RL. The role of beta interferon in the treatment of MS. Discussion Paper 96/01. Sheffield: Trent Institute for Health Services Research, Universities of Leicester, Nottingham and Sheffield; 1996.
81. Parkin D, McNamee P, Jacoby A, Miller P, Thomas S, Bates D. A cost-utility analysis on interferon-beta for multiple sclerosis. *Health Technol Assess*, 1998;2(4).
82. Rudick RA, Cohen JA, Weinstock-Guttman B, Kinkel BP, Ransohoff RM. Management of multiple sclerosis. *N Engl J Med* 1997;337:1604–10.
83. Clegg A, Bryant J, Milne R. Disease-modifying drugs for multiple sclerosis: a rapid and systematic review. *Health Technol Assess* 2000;4(9).
84. De Broe S, Christopher F, Waugh N. The role of specialist nurses in multiple sclerosis: a rapid and systematic review. *Health Technol Assess* 2001;5(17).
85. Richards RG, Sampson FC, Beard SM, Tappenden P. A review of the natural history and epidemiology of multiple sclerosis: implications for resource allocation and health economic models. *Health Technol Assess* 2002;6(10).
86. Branas P, Jordan R, Fry-Smith A, Burls A, Hyde C. Treatments for fatigue in multiple sclerosis: a rapid and systematic review. *Health Technol Assess* 2000;4(27).
87. National Institute for Clinical Excellence (NICE). Beta interferon and glatiramer acetate for the

- treatment of multiple sclerosis. Technology Appraisal Guidance No. 32. London: NICE; 2002.
88. Department of Health. Cost effective provision of disease modifying therapies for people with multiple sclerosis. Health Service Circular HSC 2002/004. Department of Health, London.
89. Rudge P. Are clinical trials of therapeutic agents for multiple sclerosis long enough? *Lancet* 1999; **353**:1033–4.
90. Compston A. Provision of treatment for multiple sclerosis. *Lancet* 1999; **353**:1710–11.
91. McNamee P, Parkin D. Cost-effectiveness of interferon beta for multiple sclerosis: the implications of new information on clinical effectiveness. *Health Technol Assess* 1998; **2**, Update October 1999.
92. Forbes RB, Lees A, Waugh N. Population based cost utility of interferon beta-1b in secondary progressive multiple sclerosis. *BMJ* 1999; **319**:1529–33.
93. Davies L, Drummond M, Papanikolaou P. Prioritizing investments in health technology assessment: can we assess the potential value for money. *Int J Technol Assess Health Care* 2000; **16**:73–91.
94. Claxton K, Neumann PJ, Araki S, Weinstein MC. Bayesian value-of-information analysis. An application to a policy model of Alzheimer's disease. *Int J Technol Assess Health Care* 2001; **17**:38–55.
95. Fryback DG, Chinnis JO, Ulvila JW. Bayesian cost-effectiveness analysis: an example using the GUSTO Trial. *Int J Technol Assess Health Care* 2001; **17**:83–97.
96. Briggs AH. A Bayesian approach to stochastic cost-effectiveness analysis: an illustration and application to blood pressure control in type 2 diabetes. *Int J Technol Assess Health Care* 2001; **17**:69–82.
97. Sheingold SH. Can Bayesian methods make data and analyses more relevant to decision makers? *Int J Technol Assess Health Care* 2001; **17**:114–22.
98. Meltzer D. Addressing uncertainty in medical cost-effectiveness analysis. Implications of expected utility maximization for methods to perform sensitivity analysis and the use of cost-effectiveness analysis to set priorities for medical research. *Journal of Health Economics* 2001; **20**:109–29.

# Appendix I

## Details of the search strategy

**G**iven the broad area and the variety of possible search terms, the aim of the search was to achieve higher sensitivity rather than specificity. Strategy used to search MEDLINE:

- 1 (priorit\$ adj3 setting).ti.
- 2 (research adj5 priorit\$).mp.
- 3 (economi\$ or cost\$1 or fund\$ or financ\$).mp.
- 4 (preliminary adj5 apprais\$).mp.
- 5 ((assess\$ or select\$ or plan\$ or evaluat\$) and (biomedical or medical or health) and (research or trial or technolog\$)).mp.
- 6 (model\$ or "decision analy\$" or scenario).mp.
- 7 "expected value of information".mp.
- 8 1 and 5
- 9 (1 and 2) not 8
- 10 4 not (8 or 9)
- 11 3 and 5
- 12 6 and 11
- 13 (priorit\$ or payback or return).mp. and 12
- 14 (priorit\$ or payback or return\$).mp.
- 15 2 and 14
- 16 7 or 8 or 9 or 10 or 13
- 17 (3 and 5 and 6) not 16
- 18 ((assess\$ or select\$) and (health or biomedical or medical) and (research or trial or technology)).mp.
- 19 (3 and 6 and 18) not 16
- 20 1 and 6
- 21 16 not 19
- 22 "value of information".mp.
- 23 18 and 22
- 24 6 and 22
- 25 2 and 3 and 6
- 26 19 and (priorit\$ or payback or return\$).mp.
- 27 algorithms/ or knowledge/ or methods/ or models, theoretical/ or research/ or clinical protocols/ or pilot projects/ or research design/
- 28 2 and 6 and 27
- 29 health care rationing/ or health care reform/ or health plan implementation/ or health planning guidelines/ or health priorities/ or

health services research/ or national health programs/ or technology assessment, biomedical/

- 30 2 and 6 and 29
- 31 27 and 6 and 2 and 3 and 29
- 32 3 and 6 and 14
- 33 Research Design/
- 34 Decision Making/
- 35 32 and (33 or 34)
- 36 20 not 35
- 37 19 and (health or research or priorit\$).mp.
- 38 35 or 36 or 37 or 7
- 39 (2 and 3) not 38

A similar search strategy was used for HealthSTAR.

The yield of the search was 707 papers, while a further 319 were identified in the updating search, which covered the period 1999 to early 2002.

- HEED

Search strategy:

1. research AND priorit\*
2. fund\* OR financ\* OR cost OR assess
3. research OR health OR biomedical
4. preliminary OR payback OR priorit\*
5. 2 AND 3 AND 4 NOT 1

In 1998 the search yielded 249 references and a further 38 references were produced in the updating search.

- BIDS IBSS:

Search all types of publications in English in the fields of Economics, Sociology and Political science.

Search for: research AND priorit\* in Title/keywords/abstract

BIDS produces 72 hits and another 97, published after 1998.



## Appendix 2

### Trial of the costs and benefits of postnatal midwifery support

A synopsis of the trial is given on the last page of this appendix.

#### **Background**

Enclosed are several papers and articles which provide information concerning the background to the research project 'The Costs and Benefits of Post-natal Midwifery Support', and to the intervention itself. The researchers have agreed for this trial to be a case study for our model which will be informed by your opinions below. As you know, we are developing a model for the NHS HTA R&D programme to assess the likely costs and benefits of undertaking specific research proposals. The model will attempt, among other things, to estimate the 'expected cost-effectiveness ratio' of a research proposal.

Within the model we will include the cost of the trial and the likely costs associated with implementation. The latter will depend upon current and likely future practice as well as the outcome of the trial itself.

#### **Likely cost of the service**

The direct cost of one Community Midwifery Support Worker (CMSW) is about £10,820 per year, of which £10,375 is salary costs (direct costs are therefore heavily dominated by the salary). The direct cost to a Health Authority of providing a comprehensive CMSW service depends upon the fertility rate in the geographical area. Using data from the 1995 Birth Statistics, the mean number of maternities in a health authority in 1995 was 6114 (SD 2454), with the median being 6087 (range 1227–15,057). If a CMSW can make three visits a day, five days a week, and each mother receives 10 daily visits within the first 28 days, on average a CMSW will visit approximately 6 mothers per month. Therefore, using a crude calculation, the direct cost to a health authority of a CMSW service will be approximately £150,000 for every 1000 births, or £150 per birth.

There may also be changes in the use of other personal health and social services by the mother and child. These may increase or decrease. Included with this paper are several journal articles relating to the use of health services by mothers in the early postpartum period. Your opinion as to the likely change in the use of health services (in particular GP visits) if any, would be welcome. In particular whether you think the utilisation of other services might increase or decrease (and if so which services and by how much).

#### **Possible trial conclusions**

In order to perform a preliminary economic evaluation to estimate the expected returns to the research, estimation is also needed of the likely benefits from the implementation of the technology. This information will come mainly from the outcome of the trial. We will ask you to consider three likely conclusions of the trial, and how Health Authorities or Trusts might implement this service accordingly. The extent of the implementation of a CMSW service would also be important. The trial evaluates a service for all mothers in the early postpartum period. However, it is possible that certain Health Authorities or providers might consider introducing the service for selected 'at risk' groups.

There are three potential conclusions to the trial:

- favourable (i.e. implementation of the experimental intervention is encouraged)
- unfavourable (i.e. implementation of the experimental intervention is discouraged)
- inconclusive (i.e. the trial fails to arrive at a significant conclusion, or there is no clear difference in cost-effectiveness).

Our questions seek your opinion regarding the likely future practice given these broad conclusions, and the costing information presented above.

## **Local postnatal service**

The level of implementation of the CMSW service is also likely to vary according to existing services, as the provision of postnatal care currently varies considerably across different localities, especially with regard to the number of midwife visits; current provision of services will affect the ease with which a Support Worker service may be introduced into an area.

It might help to consider your answers in the light of alternative levels of existing services:

### **Local policy A**

Provision of postnatal care by midwives follows approximately the national policy of not less than ten daily midwife visits.

### **Local policy B**

Selective visiting based upon the midwife's judgement regarding needs of the mother.

The following pages will make reference to these alternative local policies.

## **Questions to experts**

We should like to meet you to discuss the issues surrounding this research project. Below are the areas which we would like to discuss with you in the light of the enclosed papers, the research proposal briefing and given your own expert knowledge and experience. (Please contact us if you have questions prior to our meeting).

### **Question 1: The development of postnatal services**

Given the organisation of postnatal care around 1994 (the time of the research proposal), what structure and level of service would you have expected now? The paper by Garcia, Renfrew and Marchant (1994)<sup>49</sup> identified patterns of postnatal care in the early 1990s. Do you think this pattern would have continued, or would you have expected some reform?

### **Question 2: Implementation of the intervention following reports of the trial results**

This section will cover how postnatal services may develop following the conclusion and reporting of the trial. The areas which we wish to cover for each of the three potential conclusions of the trial, i.e.:

- Favourable (a significant difference in self-reported health status in favour of the new intervention with a favourable cost-effectiveness ratio for the CMSW service)
- Unfavourable (a significant difference in self-reported health status against the new intervention with an unfavourable cost-effectiveness ratio for the CMSW intervention)

- Inconclusive (no significant difference in self-reported health status and the cost between the control and the new intervention).
  - are:
  1. To what extent, if any, might (or would you recommend) this new service be adopted – what percentage of midwifery services might introduce a CMSW service?
  2. How might the existing local policy regarding the provision of midwifery services affect the level of implementation? (e.g. related to the alternatives local policies on page 2).
  3. If implemented, might you expect or recommend the service be provided to all women, or targeted at certain 'at risk' groups?

### **Question 3: Effects on the use of other services**

If the CMSW service were adopted, how might other relevant health services be affected, specifically:

- midwifery services (if released because the CMSW service reduced the need for midwifery visits)
- GP services.

(consider in the light of the alternative current policies given on page 2).

### **Question 4: Outcome measures used in the trial**

Do you think that the major and other outcome measures used by the researchers are the appropriate ones for the research? Which might be more likely to influence policy and/or practice (SF-36 general health perception dimension, Edinburgh Postnatal Depression Index, breast-feeding rates, SF-36 health status)? Are there alternative outcome measure(s) you think would be more appropriate?

### **Question 5: Threshold levels**

The primary end-point in the researchers' cost-effectiveness analysis is a five-point improvement in the General Health Perception dimension of the SF-36. This dimension covers five questions:

1. In general, would you say your health is
  - Excellent
  - Very good
  - Good
  - Fair
  - Poor
  
- 11a. I seem to get ill more easily than other people
  - Definitely true
  - Mostly true
  - Don't know
  - Mostly false
  - Definitely false
  
- 11c. I expect my health to get worse
  - Definitely true
  - Mostly true
  - Don't know
  - Mostly false
  - Definitely false
  
- 11b. I am as healthy as anybody I know
  - Definitely true
  - Mostly true
  - Don't know
  - Mostly false
  - Definitely false
  
- 11d. My health is excellent
  - Definitely true
  - Mostly true
  - Don't know
  - Mostly false
  - Definitely false

The trial is set to identify a five-point improvement in General Health Perception. This would be achieved by the mother moving up one score in one answer with the answers to the other questions remaining the same. For example, if the answer in 11b went from 'don't know' to 'mostly true', and all other answers remained the same, this would be approximately equivalent to a five-point improvement. This is the minimum improvement that the trial would detect. The actual improvement detected could of course be much higher.

We are interested in the level of cost-effectiveness that might trigger adoption of this intervention. Would you have some feel for how much a commissioner of postnatal services might be willing to pay to achieve an improvement? Alternatively, how much of her budget would a Director of midwifery services in a provider trust be willing to allocate per woman to achieve an improvement of so many points (e.g. 5, 15, 25 points)? Do you think there are other services that would be worth dropping to achieve this?

### Use of answers from experts

We shall run a series of preliminary economic evaluations using this information, the cost of the trial and the expected cost of implementation to determine the cost-effectiveness ratio. Your answers to the questions on the likely implementation of the intervention following the trial and the effects on other health services will be used to help estimate the expected costs and benefits from the likely implementation of the CMSW service. These preliminary economic evaluations will then be used for a 'mock report', designed to inform a research funding body on the likely returns to the research. This will be compared with the actual results when they are published, using predictors of the impact of the results.

Thank you very much

We should like to come back for your further predictions when the actual results of the trial are available.

## Postnatal care: references

- Garcia J, Renfrew M, Marchant S. Postnatal home visiting by midwives. *Midwifery* 1994;10:40-3.<sup>49</sup>
- Glazener C, Abdalla M, Russell I, Templeton A. Postnatal care: a survey of patients' experiences. *British Journal of Midwifery* 1993;1:67-74.<sup>50</sup>
- Twaddle S, Liao XH, Fylie H. An evaluation of postnatal care individualised to the needs of the woman. *Midwifery* 1993;9:154-60.<sup>51</sup>
- Van Teijlingen ER. The profession of maternity home care assistant and its significance for the Dutch midwifery profession. *Int J Nurs Stud* 1990;27:355-66.<sup>52</sup>

## Synopsis: the costs and benefits of postnatal midwifery support

### Sample

- All women delivering a live baby after 37+ weeks of gestation, birth weight 2.5 kg+. 360 women in each group.

### Intervention

- Midwifery visit on days 1 (or discharge day), 3 (or 4), 6 and 10
- Community Midwifery Support Worker (CMSW) visit daily on weekdays (10 visits)
- 24 hour telephone access to midwife until the 28th day
- CMSW will be managed within the provider Trust (Northern General Hospital), trained to NVQ level 2, and will work to protocols and guidelines for a narrow range of tasks, i.e. establish a supportive relationship, provide regular contact to recognise signs of problems, reinforce midwifery advice regarding breast feeding, etc., do light housework and encourage appropriate use of health services and other agencies.

### Control

- Current midwifery visiting practice until the 28th day
- 24 hour telephone access to midwife until the 28th day

## End-points

*Major (basis for sample size):*

- SF-36 GHP score at 6 weeks (trial size to detect five-point difference between groups, 5% significance, 80% power).

*Other end-points:*

- 6 weeks: Edinburgh Postnatal Depression Inventory, SF-36 health status, breast feeding
- 6 month follow-up: SF-36, breast feeding, maternal morbidity, acceptability of service, additional use of health and other services in first 6 months.

## Economic analysis

- To establish additional (or reduced) cost incurred by the intervention over the control, related to the benefits of the expected intervention. Cost-effectiveness in terms of cost per five-point improvement in health measure on SF-36 GHP dimension (direct and indirect costs, fixed and variable costs).

## Other information collected (at 6 weeks and 6 months)

Mother's feeling for baby, confidence, control and self-esteem, information and advice received, support from family/friends/professionals, acceptability of care.



## Appendix 3

### Trial of the costs and benefits of postnatal midwifery support: questions to experts (2)

**All answers are completely confidential**

Given the attached results of the trial of costs and benefits of the community postnatal support worker service<sup>53</sup>

1. To what extent, if any, might or would you recommend that this service now be adopted:

For all women?

For certain 'at risk' groups?

For no women?

Over what timescale do you think this change if any would be implemented?

Comments

2. Were such a service in use, as it is in certain areas, would you recommend it be:

Reduced?

Terminated?

Over what timescale do you think this change be implemented?

Comments

3. Which of the outcome measures do you think were relevant to your decision? (Mark 1, 2 ... in order of preference)

SF-36 General Health Perception dimension

Edinburgh Post Natal Depression Score

Breast-feeding rates

SF-36 health status

Client satisfaction with service

4. Are there any means of dissemination you think necessary to speed any desired change?

5. **Any overall comments on the study and/or the results**

## Appendix 4

### An RCT of infusion protocols in adult pre-hospital care

#### **Background**

Enclosed are several papers and articles which provide information concerning the background to the research project 'An RCT of Infusion Protocols in Adult Pre-Hospital Care', and to the intervention itself. The researchers have agreed for this trial to be a case study for our model, which will be informed by your opinions below. As you know, we are developing a model for the NHS HTA R&D programme to assess the likely costs and benefits of undertaking specific research proposals. The model will attempt, among other things, to estimate the 'expected incremental cost-effectiveness ratio' of a research proposal.

Within the model we will include the cost of the trial and the likely costs associated with implementation, which will depend upon current and likely future practice according to the outcome of the trial.

#### **Possible trial conclusions**

In order to perform a preliminary evaluation to estimate the expected returns to the research, estimation is needed of the likely benefits from the implementation of the technology. This information will come mainly from the outcomes of the trial, which as yet are of course unknown. We will ask you to consider three possible conclusions of the trial, and how ambulance services might implement this service accordingly. Implementation of particular protocols would also be important. The trial evaluates the benefits of 'scoop and run' compared with 'field stabilisation' for trauma patients whose estimated time of transport to hospital is between 15 and 60 minutes.

Therefore, there are three different potential conclusions to the trial:

- Favouring 'field stabilisation' for patients by using a crystalloid and colloid mixture (protocol A).
- Favouring 'scoop and run' (protocol B).

• Inconclusive (i.e. the trial fails to arrive at a significant conclusion, or there is no clear difference in cost-effectiveness between using a 'scoop and run' protocol and a 'field stabilisation' protocol). Differences in patient management between the two protocols will occur during transport to hospital, although the cost implications of this are likely to be minimal. However, differences in patient management might also occur in the A&E department, the relevant inpatient departments and in the primary and community care setting. These will produce cost differences. These differences could be due to the lower costs of patients recovering more quickly after on-scene i.v. infusion, or perhaps, greater costs due to the use of i.v. infusion saving more lives, which subsequently may increase the number of seriously disabled survivors.

## Aim

The study aims to evaluate the use of cannulation and fluid infusion at the accident scene (field stabilisation) by ambulance personnel in severely injured adults, compared with the immediate transport to hospital (scoop and run).

## Sample

Includes: adult trauma patients attended by East Anglia and Leicestershire ambulance and paramedic services, who:

- subsequently stayed in hospital for 72+ hours, or died as a result of their injuries OR
- had a triage-revised trauma score of 10 or less on scene.

Excludes: those attended by a doctor on scene, those who were transported by helicopter, those patients who died or had a revised trauma score of 0 on scene and did not show any signs of life during pre-hospital phase (poisonings, hangings, drownings and asphyxiations).

## Study design

420 paramedics will be randomised to work according to one of two protocols if the estimated time of transport of a trauma patient to hospital is between 15 and 60 minutes (see attached).

- *Protocol A*: Leans to field stabilisation (infusion by colloid/crystalloid combination if any of the indicators given in the current operational protocols for i.v. infusion are present).
- *Protocol B*: Leans to 'scoop and run' (infusion only if time to hospital is greater than 60 minutes).

It is estimated that the two ambulance services will yield over 1600 patients over one year, or 3.8 patients per paramedic.

The trial is set to detect a difference in mortality of 14% versus 20% between the two arms of the trial (significant at the 5% level and 80% power).

## Outcomes

Mortality, length of stay in hospital, length of stay in intensive care, broad disability level and general health status at two and six months after injury.

- a) Overall mortality is compared for blunt trauma patients whose estimated time of arrival at the hospital is 15–60 minutes.
- b) all adult trauma patients meeting the inclusion criteria attended by paramedics in the two main arms of the study, after standardising for case-mix (i.e. however long to hospital).

Morbidity: comparisons of disability on discharge and SF-36 scores in the two groups, adjusted for case-mix.

Subgroups: analysing outcomes by time since training and by years of experience of the paramedics.

## Economic analysis

The estimated marginal cost of operating the infusion protocol will be related to the benefits in terms of reduced fatality and/or disability. The implied marginal cost of life years saved and reduced fatality with similar implied values as obtained from other health service interventions, will be compared with the estimated marginal cost of the infusion protocol.

## Questions to experts

### Likely future practice

1. How would you have expected (in 1995) the provision of A&E pre-hospital services to have developed in the absence of this trial? In particular, would you have expected that the trend to 'field stabilisation' would have continued generally, or would this have been the case only for patients with specific injuries or conditions?

### Possible trial outcomes

2. Given the following possible outcomes of the trial:
  - Favouring scoop and run.
  - Favouring field stabilisation.
  - Inconclusive

What would you expect to be the likely change in practice? In particular, what proportion of ambulance services would you expect to introduce such a protocol?

### **Trial outcomes**

3. The outcome measures are a range of measures relating to mortality and morbidity in hospital and in the community over 6 months (mortality, length of stay in ITU, broad disability level, general health perception using SF-36 at 2 and 6 months after injury).

Do you think these are the most appropriate outcomes? Which might be more likely to influence future policy and practice in the provision of pre-hospital care to adult trauma patients? Are there any other outcomes that might have been relevant, or more relevant than those used?

### **Threshold levels**

4. The study is designed to detect a difference in mortality of 14% versus 20% between the two arms of the trial (80% power, significant at 5%). Given that there are approximately 650 adult trauma patients per annum in a non-metropolitan area covered by one ambulance service, if this minimum (6%) difference were detected, at least 41 extra patients would survive per year (about one-third fewer deaths) under the more effective protocol.

The economic evaluation will present the relative cost-effectiveness in terms of the marginal cost of following the infusion protocol related to marginal benefits in terms of reduced mortality and/or disability.

We are interested to know what barriers, if any, you think might prevent the adoption of what is found to be the more effective service.

- What sort of financial barrier might there be to the adoption?
- What sort of cost-effectiveness barrier might there be? (i.e. how big would the benefits need to be relative to the costs?)
- Might there be other barriers to adoption, e.g. organisational barriers?

We are interested to know what level of incremental costs you think would be too high for the more effective strategy to be adopted by an ambulance service, given that for an average ambulance service at least 41 more patients would be expected to survive (about one-third fewer deaths) under the more effective protocol. It might be useful to give your answer as a percentage of the annual ambulance service budget.

### **Probabilities of trial outcomes**

5. There are three potential outcomes for the trial:

- Favouring scoop and run.
- Favouring field stabilisation.
- Inconclusive.

We should like to know what you think are the most likely outcomes of the trial. Could you assign percentages to the likelihood of the trial reaching each of the three outcomes listed above, from 0% (no chance) to 100% (absolute certainty)? The sum of all three should come to 100%.

### Use of your answers

We shall run a series of preliminary economic evaluations using this information, the cost of the trial and the expected cost of implementation, to determine the cost-effectiveness ratio. Your answers to the questions on the likely practice following the trial's conclusion and the effects on other health services will be used to help estimate the expected differences in costs and benefits of an ambulance service implementing a field stabilisation protocol, compared with those of a scoop and run protocol. These preliminary economic evaluations will then be used for a 'mock report', designed to inform a research funding body on the likely returns to the research. This will be compared with the actual results when they are published, using predictors of the impact of the results.

*Thank you very much*

We should like to come back for your further predictions when the actual results of the trial are available.

## A&E fluid infusion protocol: references

### British papers

- Deakin CD, Hicks IR. AB or ABC: pre-hospital fluid management in major trauma. *Journal of Accident and Emergency* 1994;**11**:154–7.<sup>54</sup>
- Deakin C, Davies G. Defining trauma patient subpopulations for field stabilization. *Eur J Emerg Med* 1994;**1**:31–3.<sup>55</sup>
- Deakin CD. Early fluid resuscitation in haemorrhagic shock. *Eur J Emerg Med* 1994;**1**:83–5.<sup>56</sup>

### US papers for or against field stabilisation

- Aprahamian C, Thompson BM, Towne JB, Darin JC. The effect of a paramedic system on mortality of

major open intra-abdominal vascular trauma. *J Trauma* 1983;**23**:687–90.<sup>57</sup>

Pons PT, Moore EE, Cusick JM, Brunko M, Antuna B, Owens L. Pre-hospital venous access in an urban paramedic system – a prospective on-scene analysis. *J Trauma* 1988;**28**:1460–3.<sup>58</sup>

Smith JP, Bodai BI, Hill AS, Frey CF. Pre-hospital stabilization of critically injured patients: a failed concept. *J Trauma* 1985;**25**:65–70.<sup>59</sup>

### Other relevant US papers

- O'Gorman M, Trabulsky P, Pilcher DB. Zero-time pre-hospital IV. *J Trauma* 1989;**29**:84–6.<sup>60</sup>
- Wears RL, Winton CN. Load and go versus stay and play: analysis of pre-hospital IV fluid therapy by computer simulation. *Eur J Emerg Med* 1994;**1**:31–3.<sup>61</sup>

## Appendix 5

### An RCT of infusion protocols in adult pre-hospital care: questions to experts (2)

**All answers are completely confidential**

Given the attached results of the trial of infusion protocols relating to adult trauma patients whose estimated time of transport to hospital is between 15 and 60 minutes:<sup>63</sup>

1. What change in practice, if any, do you think will result from the results of the trial?

What proportion of services would now use the 'scoop and run' protocol for these patients?

What proportion of services would now use the 'field stabilisation' protocol for these patients?

2. Over what timescale do you think the change, if any, would take place?

3. What means of dissemination or other action do you think would be effective in speeding any appropriate change?

4. Which of the outcome measures were relevant for your views on practice change in the previous questions? (mark 1, 2 ... in order of importance)

Mortality

Length of stay in ITU

Broad disability level

General Health Perception using SF-36 at 2 months after injury

General Health Perception using SF-36 at 6 months after injury

5. Any overall comments on the study and/or results?

# Appendix 6

## The UK small aneurysm trial

### **Introduction**

As you know we are developing a model for the NHS HTA R&D programme to assess the likely costs and benefits of undertaking specific research proposals. The model will be used among other things, to estimate the 'expected cost-effectiveness ratio' of a research proposal.

To apply the model we will need to use the cost of the trial and the likely costs associated with implementation. The latter will depend upon current and likely future practice as well as the outcome of the trial itself.

Enclosed are several papers and articles that provide information concerning the background to the research project '*The UK Small Aneurysm Trial*'. The researchers have agreed for this trial to be a case study for our model, which will be informed by your opinions below.

### **Background to the study**

The study consists of a number of separate proposals that together form the UK Small Aneurysm Trial. Our study will consider the trial **as a whole**, and not the individual components.

This trial is designed to determine whether early elective surgery as compared with a period of ultrasound surveillance provides better management of patients with small abdominal aortic aneurysms (4.0–5.5 cm in diameter). The study is a multicentre randomised trial, including approximately 1000 patients aged 60–76 years, randomised to either surgery or ultrasonographic surveillance.

The study will compare long-term survival over 5 years and the costs between the surgery and surveillance groups. Health-related quality of life will also be compared between the two groups.

### **Aim**

The study aims to compare the mortality, quality of life and health-service costs associated with early surgical repair compared with surveillance in the management of small abdominal aortic aneurysms.

### **Sample and study design**

Patients aged 60–76 years with asymptomatic abdominal aortic aneurysms (4–5.5 cm) will be recruited from referrals to the participating vascular surgeons and coordinated from five centres. More than 120 vascular surgeons will participate in the trial and 1000 patients will be

recruited, 500 in the first year and 250 in each successive year of the project. Patients will be followed up for an average of 5 years each, making 5000 patient years of follow-up.

### **Outcomes**

Mortality: all-cause mortality at 2, 4 and 6 years will be compared between the surgical repair and surveillance groups.

Quality of life: measured by the Medical Outcomes Study short-form general health survey prior to randomisation, and every 3 months or 6 months (depending on size of aneurysm) for the duration of observation, or 1 year after surgical repair.

## **Economic analysis**

The costs to the NHS of the two alternative management protocols will be compared, and used in conjunction with the mortality results and the quality of life comparisons to make recommendations on the management of small asymptomatic abdominal aortic aneurysms.

### **Possible trial conclusions**

We should like to ask you to consider three possible conclusions of the trial, and make an informed guess at how NHS Trusts, relevant departments or consultants, might, for each conclusion, adjust their treatment of patients with small abdominal aortic aneurysms.

There are three potential conclusions to the trial:

- Favouring surgery (improved mortality for surgery compared with surveillance, favourable balance with quality of life and cost-effectiveness, significant difference in cost-effectiveness and improved quality of life in comparison to patients under surveillance).
- Favouring surveillance of patients (no significant improvement in mortality after surgery, favourable balance of quality of life between surgery and surveillance and a preferable cost-effectiveness ratio for surveillance).
- Inconclusive between surgery and surveillance (where neither of the scenarios above dominates, e.g. no clear difference in the cost-effectiveness ratio, taking into account mortality, quality of life and costs).

We would like your opinion regarding future practice under each of the three scenarios (see over).

## **Questions to experts**

### **Likely future practice**

1. To assess the effect of the proposed trial on practice, we need to have some idea of how practice would have developed in the absence of this trial. The initial proposal was submitted in 1990. How would you have expected (in 1990) the treatment of patients aged 60–76 with small abdominal aortic aneurysms to have developed? What proportion of patients with small aneurysms would have undergone early elective surgery in 1990, and how would this have developed over the subsequent 10 years if this trial had not proceeded?

### **Possible trial outcomes and implementation**

2. Given each of the following possible scenario outcomes of the trial, as defined on the previous page:

- Favouring surgery
- Favouring surveillance
- Inconclusive between surgery and surveillance

What would you expect to be the likely development of management following each of these conclusions? In particular, what proportion of cardiovascular surgeons or NHS trusts would employ the more effective and/or cost-effective management procedure?

How might practice change if the trial reached the ‘inconclusive’ outcomes?

### Trial outcomes

3. The primary outcome measure used in the trial is all cause mortality over 5 years. A secondary measure is health-related quality life as measured by the Medical Outcomes Study short-form general health survey. Quality of life is measured prior to randomisation and every 3 months or 6 months (depending on the size of the aneurysm) for the duration of ultrasonographic surveillance, or 1 year after aneurysm repair. Health-service costs were also measured as part of the trial outcome.

Do you think that these are the most appropriate outcome measures? Are there any other outcomes that might have been relevant, or more relevant than those used?

### Threshold levels

4. The study is designed to detect a 5 year difference in mortality of at least 9% with 80% power, significant at 5% between the two arms of the trial (e.g. a 5 year mortality associated with surgery of 29% with 38% in the surveillance group). In addition, surgical repair may be associated with a higher level of quality of life. What combination of improved mortality and improved quality of life and increased costs would be acceptable to surgeons or NHS trusts to trigger a change in practice?

We are interested to know what barriers and threshold levels might exist concerning the adoption of the more cost-effective management protocol:

- What reduction in mortality due to surgical repair would be sufficient to persuade vascular surgeons to change from observation to early surgical repair?
- What sort of financial barrier might there be to the adoption? Specifically, if the minimum detectable reduction in mortality of 9% due to surgical repair were found to be significant, what level of net incremental cost associated with surgery might prevent the adoption?
- What level of improvement in quality of life (percentage improvement in the Medical Outcomes Study short-form general health survey score), if any, would be sufficient to encourage the early surgical repair of aneurysms, if the incremental cost per patient was:
  - (1) £100
  - (2) £500
  - (3) £1000
  - (4) £1500
  - (5) £2000?

### Probabilities of trial outcomes

5. There are three potential outcomes for the trial (defined on page 1):

- In favour of surgical repair.
- In favour of ultrasonographic surveillance.
- Inconclusive between surgery and surveillance.

What do you think is the most likely outcome of the trial? Could you assign percentages to the likelihood of the trial reaching each of the three outcomes listed above, from 0% (no chance) to 100% (absolute certainty)? The sum of all three should come to 100%.

**Implications for national screening programme**

6. What possible implications, if any, do you think the possible different outcomes would have for a national screening programme?

**Use of expert answers**

We shall run a series of preliminary economic evaluations using the cost of the trial and the expected cost of changed management to determine the cost-effectiveness ratio. Your answers to the questions on the likely changed management or otherwise for each scenario outcome of the trial will be used to help estimate the expected costs and benefits from the trial. These preliminary economic evaluations will then be used for a 'mock report', designed to inform a research funding body on the likely returns to the research. This will be compared with the actual results when they are published, using predictors of the impact of the results.

Thank you very much

We should like to come for your further predictions when the actual results of the trial are available.

**UK small aneurysm trial:  
references**

Bengtsson H, Bergqvist D, Jendteg S, Lindgren B, Perrson U. Ultrasonographic screening for abdominal aortic aneurysm: analysis of surgical decisions for cost-effectiveness. *World J Surg* 1989;12:266–71.<sup>64</sup>

Cheatle TR, Scull JH. Abdominal aortic aneurysms: a review of current problems. *Br J Surg* 1989;76:826–9.<sup>65</sup>

Collin J, Araujo L, Walton J, Lindsell D. Oxford screening programme for abdominal aortic aneurysm in men aged 65 to 74 years. *Lancet* 1998;ii:613–15.<sup>66</sup>

Quill DS, Colgan MP, Summer DS. Ultrasonic screening for the detection of abdominal aortic aneurysms. *Surg Clin North Am* 1989;69:713–20.<sup>67</sup>

## Appendix 7

### The UK small aneurysm trial: questions to experts (2)

**All answers are completely confidential**

Given the attached results<sup>68,69</sup> of the UK small aneurysm (size 4.0–5.5 cm for patients aged 60–76) trial:

1. In your judgement, what proportion of cardiovascular surgeons would now treat such aneurysms

With surgery?

With ultrasonic surveillance?

2. Do you think there would be likely difference in choice of treatment within these ranges according to age or size of aneurysm?

#### Detail

3. Which of the outcome measures were relevant to your judgement? (Mark 1, 2 ... in order of importance)

All cause mortality over 5 years

Quality of life on Medical Outcome Study Health Survey

Health services costs

4. Do you think these results may change treatment for patients outside the ranges of the study?
  5. What implications do you think the results have for a possible national screening programme?
  6. What means of dissemination of the results or other action do you think is necessary for appropriate implementation?
  7. Any overall comments on the study and/or the results?

THANK YOU VERY MUCH



# Health Technology Assessment Programme

## Prioritisation Strategy Group

### **Members**

<b>Chair,</b> <b>Professor Kent Woods,</b> Director, NHS HTA Programme & Professor of Therapeutics, University of Leicester	Professor Bruce Campbell, Consultant Vascular & General Surgeon, Royal Devon & Exeter Hospital	Dr John Reynolds, Clinical Director, Acute General Medicine SDU, Radcliffe Hospital, Oxford
	Professor Shah Ebrahim, Professor in Epidemiology of Ageing, University of Bristol	Dr Ron Zimmern, Director, Public Health Genetics Unit, Strangeways Research Laboratories, Cambridge

## HTA Commissioning Board

### **Members**

<b>Programme Director,</b> Professor Kent Woods, Director, NHS HTA Programme, Department of Medicine and Therapeutics, Leicester Royal Infirmary, Robert Kilpatrick Clinical Sciences Building, Leicester	Professor John Brazier, Director of Health Economics, Sheffield Health Economics Group, School of Health & Related Research, University of Sheffield, Scharr Regent Court, Sheffield	Professor Adrian Grant, Director, Health Services Research Unit, University of Aberdeen, Drew Kay Wing, Polwarth Building, Foresterhill, Aberdeen	Professor David Neal, Professor of Surgical Oncology, Oncology Centre, Addenbrooke's Hospital, Cambridge
<b>Chair,</b> <b>Professor Shah Ebrahim,</b> Professor in Epidemiology of Ageing, Department of Social Medicine, University of Bristol, Canyng Hall, Whiteladies Road, Bristol	Dr Andrew Briggs, Public Health Career Scientist, Health Economics Research Centre, University of Oxford, Institute of Health Sciences, Oxford	Professor Alastair Gray, Director, Health Economics Research Centre, University of Oxford, Institute of Health Sciences, Headington, Oxford	Professor Tim Peters, Professor of Primary Care Health Services Research, Division of Primary Health Care, University of Bristol, Cotham House, Cotham Hill, Bristol
<b>Deputy Chair,</b> <b>Professor Jenny Hewison,</b> Professor of Health Care Psychology, Academic Unit of Psychiatry and Behavioural Sciences, University of Leeds School of Medicine, Leeds	Dr Christine Clark, Medical Writer & Consultant Pharmacist, Cloudside, Rossendale, Lancs <i>and</i> Principal Research Fellow, Clinical Therapeutics in the School of Pharmacy, Bradford University, Bradford	Professor Mark Haggard, Director, MRC ESS Team, CBU Elsworth House, Addenbrooke's Hospital, Cambridge	Professor Ian Roberts, Professor of Epidemiology & Public Health, Intervention Research Unit, London School of Hygiene and Tropical Medicine, London
Professor Douglas Altman, Professor of Statistics in Medicine, Centre for Statistics in Medicine, Oxford University, Institute of Health Sciences, Cancer Research UK Medical Statistics Group, Headington, Oxford	Professor Nicky Cullum, Director of Centre for Evidence Based Nursing, Department of Health Sciences, University of York, Research Section, Seebom Rowntree Building, Heslington, York	Professor F D Richard Hobbs, Professor of Primary Care & General Practice, Department of Primary Care & General Practice, University of Birmingham, Primary Care and Clinical Sciences Building, Edgbaston, Birmingham	Professor Peter Sandercock, Professor of Medical Neurology, Department of Clinical Neurosciences, University of Edinburgh, Western General Hospital NHS Trust, Bramwell Dott Building, Edinburgh
Professor John Bond, Professor of Health Services Research, Centre for Health Services Research, University of Newcastle, School of Health Sciences, Newcastle upon Tyne	Dr Andrew Farmer, Senior Lecturer in General Practice, Department of Primary Health Care, University of Oxford, Institute of Health Sciences, Headington, Oxford	Professor Peter Jones, Head of Department, University Department of Psychiatry, University of Cambridge, Addenbrooke's Hospital, Cambridge	Professor Martin Severs, Professor in Elderly Health Care, Portsmouth Institute of Medicine, Health & Social Care, St George's Building, Portsmouth
	Professor Fiona J Gilbert, Professor of Radiology, Department of Radiology, University of Aberdeen, Lilian Sutton Building, Foresterhill, Aberdeen	Professor Sallie Lamb, Research Professor in Physiotherapy/Co- Director, Interdisciplinary Research Centre in Health, Coventry University, Coventry	Dr Jonathan Shapiro, Senior Fellow, Health Services Management Centre, Park House, Birmingham
		Dr Donna Lamping, Senior Lecturer, Health Services Research Unit, Public Health and Policy, London School of Hygiene and Tropical Medicine, London	

Current and past membership details of all HTA 'committees' are available from the HTA website ([www.ncchta.org](http://www.ncchta.org))

## Diagnostic Technologies & Screening Panel

### **Members**

<b>Chair,</b> <b>Dr Ron Zimmern</b> , Director of the Public Health Genetics Unit, Strangeways Research Laboratories, Cambridge	Dr David Elliman, Consultant in Community Child Health, London	Mr Tam Fry, Honorary Chairman, Child Growth Foundation, London	Mr Tony Tester, Chief Officer, South Bedfordshire Community Health Council, Luton
Dr Paul Cockcroft, Consultant Medical Microbiologist/ Laboratory Director, Public Health Laboratory, St Mary's Hospital, Portsmouth	Dr Andrew Farmer, Senior Lecturer in General Practice, Institute of Health Sciences, University of Oxford	Dr Susanne M Ludgate, Medical Director, Medical Devices Agency, London	Dr Andrew Walker, Senior Lecturer in Health Economics, University of Glasgow
Professor Adrian K Dixon, Professor of Radiology, Addenbrooke's Hospital, Cambridge	Dr Karen N Foster, Clinical Lecturer, Dept of General Practice & Primary Care, University of Aberdeen	Dr William Rosenberg, Senior Lecturer and Consultant in Medicine, University of Southampton	Professor Martin J Whittle, Head of Division of Reproductive & Child Health, University of Birmingham
	Professor Jane Franklyn, Professor of Medicine, University of Birmingham	Dr Susan Schonfield, CPHM Specialised Services Commissioning, Croydon Primary Care Trust	Dr Dennis Wright, Consultant Biochemist & Clinical Director, Pathology & The Kennedy Galton Centre, Northwick Park & St Mark's Hospitals, Harrow
	Professor Antony J Franks, Deputy Medical Director, The Leeds Teaching Hospitals NHS Trust	Dr Margaret Somerville, Director of Public Health, Teignbridge Primary Care Trust, Devon	

## Pharmaceuticals Panel

### **Members**

<b>Chair,</b> <b>Dr John Reynolds</b> , Clinical Director, Acute General Medicine SDU, Oxford Radcliffe Hospital	Dr Christopher Cates, GP and Cochrane Editor, Bushey Health Centre, Bushey, Herts.	Mrs Sharon Hart, Managing Editor, <i>Drug &amp; Therapeutics Bulletin</i> , London	Dr Ken Stein, Senior Lecturer in Public Health, University of Exeter
Professor Tony Avery, Professor of Primary Health Care, University of Nottingham	Mr Charles Dobson, Special Projects Adviser, Department of Health	Dr Christine Hine, Consultant in Public Health Medicine, Bristol South & West Primary Care Trust	Professor Terence Stephenson, Professor of Child Health, University of Nottingham
Professor Iain T Cameron, Professor of Obstetrics & Gynaecology, University of Southampton	Dr Robin Ferner, Consultant Physician and Director, West Midlands Centre for Adverse Drug Reactions, City Hospital NHS Trust, Birmingham	Professor Robert Peveler, Professor of Liaison Psychiatry, Royal South Hants Hospital, Southampton	Dr Richard Tiner, Medical Director, Association of the British Pharmaceutical Industry, London
Mr Peter Cardy, Chief Executive, Macmillan Cancer Relief, London	Dr Karen A Fitzgerald, Pharmaceutical Adviser, Bro Taf Health Authority, Cardiff	Dr Frances Rotblat, CPMP Delegate, Medicines Control Agency, London	Professor Dame Jenifer Wilson-Barnett, Head of Florence Nightingale School of Nursing & Midwifery, King's College, London
	Professor Alastair Gray, Professor of Health Economics, Institute of Health Sciences, University of Oxford	Mrs Katrina Simister, New Products Manager, National Prescribing Centre, Liverpool	

## Therapeutic Procedures Panel

### Members

**Chair,**  
**Professor Bruce Campbell,**  
Consultant Vascular and  
General Surgeon, Royal Devon  
& Exeter Hospital

Dr Mahmood Adil, Head of  
Clinical Support & Health  
Protection, Directorate of  
Health and Social Care (North),  
Department of Health,  
Manchester

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

Mr Michael Clancy, Consultant  
in A & E Medicine,  
Southampton General Hospital

Dr Carl E Counsell, Senior  
Lecturer in Neurology,  
University of Aberdeen

Dr Keith Dodd, Consultant  
Paediatrician, Derbyshire  
Children's Hospital, Derby

Professor Gene Feder, Professor  
of Primary Care R&D, Barts &  
the London, Queen Mary's  
School of Medicine and  
Dentistry, University of London

Ms Bec Hanley, Freelance  
Consumer Advocate,  
Hurstpierpoint, West Sussex

Professor Alan Horwich,  
Director of Clinical R&D, The  
Institute of Cancer Research,  
London

Dr Phillip Leech, Principal  
Medical Officer for Primary  
Care, Department of Health,  
London

Mr George Levy, Chief  
Executive, Motor Neurone  
Disease Association,  
Northampton

Professor James Lindesay,  
Professor of Psychiatry for the  
Elderly, University of Leicester

Dr Mike McGovern, Senior  
Medical Officer, Heart Team,  
Department of Health, London

Dr John C Poundsford,  
Consultant Physician, North  
Bristol NHS Trust

Professor Mark Sculpher,  
Professor of Health Economics,  
Institute for Research in the  
Social Services, University of  
York

Dr L David Smith, Consultant  
Cardiologist, Royal Devon &  
Exeter Hospital

Professor Norman Waugh,  
Professor of Public Health,  
University of Aberdeen

## Expert Advisory Network

### Members

Mr Gordon Aylward, Chief Executive, Association of British Health-Care Industries, London	Professor Pam Enderby, Professor of Community Rehabilitation, Institute of General Practice and Primary Care, University of Sheffield	Professor Allen Hutchinson, Director of Public Health & Deputy Dean of SCHARR, Department of Public Health, University of Sheffield	Professor Jon Nicholl, Director of Medical Care Research Unit, School of Health and Related Research, University of Sheffield
Ms Judith Brodie, Head of Cancer Support Service, Cancer BACUP, London	Mr Leonard R Fenwick, Chief Executive, Newcastle upon Tyne Hospitals NHS Trust	Professor Rajan Madhok, Medical Director & Director of Public Health, Directorate of Clinical Strategy & Public Health, North & East Yorkshire & Northern Lincolnshire Health Authority, York	Mrs Julietta Patnick, National Co-ordinator, NHS Cancer Screening Programmes, Sheffield
Mr Shaun Brogan, Chief Executive, Ridgeway Primary Care Group, Aylesbury, Bucks	Professor David Field, Professor of Neonatal Medicine, Child Health, The Leicester Royal Infirmary NHS Trust	Professor David Mant, Professor of General Practice, Department of Primary Care, University of Oxford	Professor Chris Price, Visiting Chair – Oxford, Clinical Research, Bayer Diagnostics Europe, Cirencester
Ms Tracy Bury, Project Manager, World Confederation for Physical Therapy, London	Mrs Gillian Fletcher, Antenatal Teacher & Tutor and President, National Childbirth Trust, Henfield, West Sussex	Professor Alexander Markham, Director, Molecular Medicine Unit, St James's University Hospital, Leeds	Ms Marianne Rigge, Director, College of Health, London
Mr John A Cairns, Professor of Health Economics, Health Economics Research Unit, University of Aberdeen	Ms Grace Gibbs, Deputy Chief Executive, Director for Nursing, Midwifery & Clinical Support Servs., West Middlesex University Hospital, Isleworth, Middlesex	Dr Chris McCall, General Practitioner, The Hadleigh Practice, Castle Mullen, Dorset	Professor Sarah Stewart-Brown, Director HSRU/Honorary Consultant in PH Medicine, Department of Public Health, University of Oxford
Professor Howard Stephen Cuckle, Professor of Reproductive Epidemiology, Department of Paediatrics, Obstetrics & Gynaecology, University of Leeds	Dr Neville Goodman, Consultant Anaesthetist, Southmead Hospital, Bristol	Professor Alistair McGuire, Professor of Health Economics, London School of Economics	Professor Ala Szczepura, Professor of Health Service Research, Centre for Health Services Studies, University of Warwick
Professor Nicky Cullum, Director of Centre for Evidence Based Nursing, University of York	Professor Robert E Hawkins, CRC Professor and Director of Medical Oncology, Christie CRC Research Centre, Christie Hospital NHS Trust, Manchester	Dr Peter Moore, Freelance Science Writer, Ashtead, Surrey	Dr Ross Taylor, Senior Lecturer, Department of General Practice and Primary Care, University of Aberdeen
Dr Katherine Darton, Information Unit, MIND – The Mental Health Charity, London	Professor F D Richard Hobbs, Professor of Primary Care & General Practice, Department of Primary Care & General Practice, University of Birmingham	Dr Andrew Mortimore, Consultant in Public Health Medicine, Southampton City Primary Care Trust	Mrs Joan Webster, Consumer member, HTA – Expert Advisory Network
Professor Carol Dezateux, Professor of Paediatric Epidemiology, London	Dr Sue Moss, Associate Director, Cancer Screening Evaluation Unit, Institute of Cancer Research, Sutton, Surrey		
Professor Martin Eccles, Professor of Clinical Effectiveness, Centre for Health Services Research, University of Newcastle upon Tyne			



## **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.***