

**Assessing the
Payback from Health
R&D:
From *ad hoc* Studies
to Regular Monitoring**

Martin Buxton
Bronwyn Croxson
Steve Hanney

Institute of Child Health
and Great Ormond Street Hospital for Children NHS Trust



Assessing the Payback from Health R&D: From *ad hoc* studies to regular monitoring

Martin Buxton¹
Bronwyn Croxson²
Steve Hanney¹

¹ Health Economics Research Group, Brunel University

² R&D Office, Institute of Child Health/Great Ormond Street Hospital, London

HERG Research Report No.27

December 1999

Price £5.00

ISBN 1-902316-10-X

HERG

HEALTH ECONOMICS RESEARCH GROUP

BRUNEL UNIVERSITY, UXBRIDGE, MIDDLESEX UB8 3PH, UK

ABOUT HERG

HERG is a research group within the Department of Economics and Finance at Brunel University. The focus of its work is the methods and practice of economic evaluation of health care interventions and programmes. The group is funded by a programme grant from the Department of Health, as well as by a number of project-related grants from funders including the MRC and the NHS Executive.

HERG DISCUSSION PAPERS AND RESEARCH REPORTS

HERG publishes a series of discussion papers to report on work in progress, and a series of research reports detailing completed research activity. Comments are encouraged and should be directed to the authors.

FURTHER INFORMATION

Information about HERG can be obtained from:

Health Economics Research Group
Brunel University
Uxbridge
Middlesex UB8 3PH
UK

Phone: 01895 274000
Fax: 01895 203330
E-mail: herg@brunel.ac.uk
Web home page: www.brunel.ac.uk/depts/herg

CONTENTS

Preface	i
Executive summary	ii
Chapter 1 : Introduction	1
Chapter 2 : Review of existing approaches to assessing the payback from research	4
Introduction	4
Categories of research payback	5
General issues in the measurement of payback	7
Methods of measuring payback	9
Conclusions	14
Chapter 3 : Characteristics of a routine monitoring system	15
Introduction	15
Assessment approaches and characteristics of systems	15
General characteristics of performance measurement systems and their relevance to research	17
Criteria for a routine monitoring system for R&D	20
Conclusions	22
Chapter 4 : Differences between research types	23
'Basic' versus 'applied' research	23
Social care research	26
Commissioned versus responsive R&D	27
Conclusions	28
Chapter 5 : What units of research?	29
Introduction	29
The main streams of research and the levels within them	29
Applying indicators to the three streams	30
Conclusions	35
Chapter 6 : Proposed comprehensive routine monitoring system	36
Sources of information for a routine monitoring system	37
Outputs and outcomes to be measured	39
Key issues	42
Conclusions	48
Chapter 7 : Research and monitoring	49
Issues for feasibility testing	49
Issues for specific evaluation once a system is implemented	49
Research opportunities provided by the database	50
Other important related research issues	50
Value in use	50
Final comments	51
References	52

List of Tables

Table 1: Categories of research payback.....	6
Table 2: Characteristics of different approaches to assessing payback	10
Table 3: Outputs and outcomes by measurement vehicle	40
Table 4: Financial attribution of outputs	44

List of Figures

Figure 1: Streams and levels of funding within an agency	31
--	----

Preface

This report was commissioned by the R&D Directorate, of the Department of Health, as a background document for an international workshop held in November 1999. The workshop was designed to discuss issues related to developing a system for monitoring payback from health-related R&D and to provide the Department with advice. It brought together civil servants, NHS managers and R&D directors, members of the research community, representatives of the research councils and other major research funding bodies. It was funded by the Department of Health, with sponsorship from the Wellcome Trust. This published report is a slightly amended version of the paper presented at the workshop and as such it does not attempt to reflect the very interesting and useful discussions held at the workshop.

We gratefully acknowledge help we received in preparing this report, through the comments of Jonathan Grant, Barbara Byth, Justin Keen, Tim Packwood, and members of the R&D Directorate, Department of Health, and through the efficient secretarial assistance from Nicky Dunne and Karen Arnold. All of the views contained in this document are those of the authors alone.

EXECUTIVE SUMMARY

Chapter 1 : Introduction

- The increasing demands for the benefits of payback from publicly funded R&D to be assessed are based partly on the need to justify or account for expenditure on R&D, and partly on the desire for information to assist resource allocation and the better management of R&D funds. The former consideration is particularly strong in relation to the R&D expenditure that comes out of the wider NHS budget.
- In this report a range of categories of payback will be identified along with a variety of methods for assessing them.
- The aim of the report is to make recommendations as to how the outcomes from health research might best be monitored on a regular basis. The specific context of the report is the NHS R&D Programme but many of the issues will be relevant for a wide range of funders of health R&D.
- The introduction sets out not only a plan of the report but also suggests that readers familiar with the general arguments and existing literature may choose to jump to Chapter 6.

Chapter 2 : Review of Existing Approaches to Assessing the Payback from Research

- Existing work describes various approaches to valuing research. Some are *ex ante* and attempt to predict the outcomes of research being considered, others are *ex post* or retrospective.
- The five categories of benefit or payback from health R&D that have been identified involve contributions: to knowledge; to research capacity and future research; to improved information for decision making; to the efficiency, efficacy and equity of health care services; and to the nation's economic performance. These are shown in Table 1 of the report
- The process by which R&D generates final outcomes can be modelled as a sequence. This includes primary outputs such as publications; secondary outputs in the form of policy or administrative decisions; and final outcomes which comprise the health and economic benefits. Feedback loops are also introduced and mitigate the limitations of a linear approach.
- Qualitative and quantitative approaches can be used but there are immense problems with time lags and attributing outcomes, and sometimes even outputs, to specific items of research funding.

- Four common methods of measuring payback can be used. Expert review, by peers or, sometimes, users is the traditional way of assessing the quality of research. Bibliometric techniques can involve not only counting publications but also using datasets such as the Science Citation Index and Wellcome's Research Outputs Database (ROD). The various methods of economic analysis of payback are difficult to undertake given the costs and problems of acquiring relevant information and estimating benefits. Social science methods include case studies, which can provide useful information but are resource intensive, and questionnaires to researchers and potential research users.

Chapter 3 : Characteristics of a Routine Monitoring System

- In moving from *ad hoc* or research studies of payback towards a more regular monitoring it is noted that whereas there has always been a tradition of evaluation of research, in the public services in general there is now a greater emphasis on audit and performance measurement and indicators. A review of these various systems suggests we should be looking to develop a system of outcomes monitoring that incorporates performance indicators (PIs) and measurement rather than an audit system that is trying to monitor activities against predetermined targets.
- Standard characteristics of performance measurement systems do not necessarily apply to research where, for example, there are non-standard outputs. Difficulties have arisen in the USA in attempting to apply the Government Performance and Results Act to research funding agencies. It is shown that because the findings of basic research, in particular, enter a knowledge pool in which people and ideas interact, it is difficult to use a PIs' approach to track eventual outcomes. However, for some types of health research it has proved more feasible to trace the flow between research outputs and outcomes.
- An outcomes monitoring system could be useful if it met the following criteria: relevant to, with as comprehensive coverage as possible of, the funders objectives; relevant to the funder's decision making processes; encourages accurate compliance; minimises unintended consequences; and has acceptable costs.

Chapter 4 : Differences Between Research Types

- The range of differences between types of research can be relevant for the design of a routine monitoring system. The OECD distinguishes between basic research, applied research and experimental development. Most DH/NHS research is applied. There

might be more of a tradition of publication of findings in applied research in health than in other fields. Nevertheless, the publication and incentives patterns operating in basic research mean that it would be inappropriate to use bibliometric indicators in a simple way across all fields even in health research.

- Despite having some differences from health research in publication patterns and in the detailed categories of payback, the broad approach proposed in Chapter 6 could be applied to social care research.
- Research that is commissioned, especially by the government, has some of the minimum conditions built into it that are associated with outcomes being generated, in particular because the funder has identified that a contribution in this area will be valuable.

Chapter 5 : What Units of Research?

- The term programme has various meanings including being used to describe a collection of projects on a common theme and to describe a block of funding for a research unit.
- Three main streams or modes of funding can be identified: projects, which are administratively grouped into programmes including a responsive programme; institutions/centres/units; individual researchers. These 3 streams are displayed in Figure 1. It is probable that the regular data-gathering for a monitoring system would operate at the basic level of each stream or mode.
- Previous work demonstrates that the full range of benefits can sometimes be applied at the level of projects, either in the responsive mode or in programmes, through the use of questionnaires to researchers. Expert and user review and user surveys have also been applied.
- Institutions and centres increasingly have experience not only of traditional periodic expert review but also of producing annual reports, although there are debates about what dimensions to include in such reviews and reports.
- Individuals in receipt of research development awards have completed questionnaires during and after the awards. These concentrate on the development of research capacity but can go wider.

Chapter 6 : A Possible Comprehensive Outcomes Monitoring System

- The proposed system is intended for DH/NHS to monitor the outcomes from its R&D in order to justify the R&D expenditure and assist with managing the portfolio. More detailed information is required for the latter purpose.

- We propose a multidimensional approach be adopted to cover all the dimensions of payback and that information be gathered from three sets of sources and Table 3 shows which methods would cover which output/outcome categories.
- Firstly, possibly annually, a questionnaire (possibly electronic) covering most payback categories should gather data from the basic level of each funding stream ie. from lead researchers of projects, from research institutions/centres, and from individual award holders.
- Secondly, supplementary information should be gathered from external databases (including the citation indices and Wellcome's ROD).
- Thirdly, a range of approaches ie. user surveys, reviews by experts and peers, case studies including economic evaluations, and analysis of sources used in policy documents such as NICE guidelines, would be undertaken on a sample basis. They would provide not only supplementary information but, as with the external databases, would also verify the data collected directly from researchers.
- These proposals can be evaluated against the criteria set out in Chapter 3:
 - The system is relevant to DH's objectives of generating payback in a range of categories.
 - Various problems have to be overcome before the system could be fully decision relevant. Firstly it might be necessary to ask researchers to apportion the contribution made to specific outputs from various funding streams. Second, to be decision relevant the information would have to be analysed and presented in a manner consistent with funders' decision making processes. This would involve a) showing how for each outcome and output, for example publications, data from one project or stream could be compared with those from another and b) demonstrating how different outputs and outcomes could be aggregated.
 - The questions of accuracy of data, minimisation of unintended consequences and the acceptability of the net costs are also addressed.

Chapter 7 : Research and Monitoring

- Whilst this report is primarily concerned with moving from *ad hoc* studies towards a routine monitoring system there are issues that need further research.
- Before embarking on full implementation the feasibility needs to be tested of items such as on-line recording of data and asking researchers to attribute proportions of research outputs to separate funding agencies.

- Once the system is implemented the value of some items can be better assessed, for example the additional value provided by self reporting of publications beyond that gained from relying on external databases.
- The data provided by the system would provide opportunities for further payback research on, for example, links between publications and other categories of payback.
- Some items such as network analysis could potentially be added to the monitoring system after further examination of them.
- Finally the benefit from the monitoring system itself should be assessed.

CHAPTER 1 : INTRODUCTION

In a number of countries, over the past decade, the evaluation of the returns from public sector investment in research and development (R&D) has moved up the policy agenda. The desire for evaluation is partly a response to demands for increased accountability to taxpayers for public expenditure; partly a result of a desire to improve the information base for decisions about allocating research funds; and partly a means of gaining information to enable improved management of research. In the US, for example, the 1993 Government Performance and Results Act has made evaluation a formal requirement for all government-supported agencies, including the National Institutes of Health (NIH, 1998).

In the UK, where public funding for health research comes through a number of channels, there is increasing pressure for systematic evaluation of health related R&D. In particular changes to the system for allocating the resources within the NHS Research and Development Programme budget have meant that attention has to be directed to evaluation (Black, 1997). The direct opportunity cost of this substantial investment of NHS funding is spending on patient care. NHS research is expected to generate returns, through improvements in health and welfare, and thus there is pressure to justify the size of the total allocation of resources to research, as opposed to health services. The so-called 'Culyer Reforms', for example, mean that a portion of the NHS R&D budget is directly contestable, requiring that methods be developed for determining the benefits likely to accrue from alternative uses of funds and for monitoring their use.

Complementing the policy-related attention, there is a growing body of academic literature developing the theory and methodology of research evaluation, and undertaking practical studies. One stream of this work has identified a range of benefits, often called 'paybacks' or 'impacts', from health-related research. These include benefits in the form of contributions to knowledge; to increased research capacity; to improved information for decisions; to the efficacy, efficiency and equity of health care services; and to economic performance. A variety of methods have been used to establish and measure these categories of benefit, including analysis of publications, peer review, questionnaires to researchers, case studies, and cost-effectiveness analyses.

Most of the work undertaken to date in the UK on payback from health research has itself been 'research' rather than 'audit', and it has not addressed the design of systematic, routine methods for evaluating benefits from research activities on an ongoing basis. A key issue is

how far it is feasible to build on existing work and develop a system suitable for widespread implementation.

It is this issue that this report addresses, with the aim of making recommendations as to how the outcomes from health research might better be monitored on a regular basis. It therefore refers selectively to existing literature rather than attempting a review of the full range of available literature. Whilst the specific context for these recommendations is the NHS R&D Programme, many of the issues and the main elements of the proposals are potentially relevant to a range of other funders of health research both in the UK and overseas. Indeed to succeed, some of the proposals may need the tacit or even active support of co-funders of research partially supported by the NHS. Moreover, it might be desirable to adopt a co-ordinated system, to reduce the administrative burden on researchers and minimise the scope for strategic behaviour. Thus it is hoped that the ideas presented here will serve not only as a basis for discussion within the NHS Programme but also more widely amongst other research funding organisations.

The early chapters of the report provide background information about issues relevant to measuring the value of R&D. Chapter 2 describes the various dimensions that contribute to payback from R&D and outlines alternative ways that payback can be measured. Chapter 3 describes the characteristics of routine monitoring systems in general, and how these might be adapted to suit the context of health related R&D. Chapter 4 discusses the way in which differences between the various types of R&D might affect the design of a routine monitoring system. Chapter 5 discusses the various streams of research and different levels at which research can be monitored. Readers familiar with the general arguments and existing literature may choose to omit the earlier chapters. The key proposals are presented in Chapter 6, which suggests a system for routinely monitoring R&D funded by the Department of Health and NHS. The final chapter, Chapter 7, suggests a number of areas for further research, some of which are necessary to check the feasibility of the system sketched in Chapter 6, before it can be operationalised. Others are necessary to underpin, test and develop the methodology after it has been implemented.

CHAPTER 2 : REVIEW OF EXISTING APPROACHES TO ASSESSING THE PAYBACK FROM RESEARCH

Introduction

There is an existing body of work addressing the value of research (OST and PREST, 1993; Salter and Martin, 1999). Some analysis seeks to inform decisions about the level of funding that should be allocated to research and how it should be allocated between alternative projects, programmes or institutions. Some studies have been described as '*ex ante*' or 'prospective', in other words seeking to predict the outcomes of research which is being considered or planned. Other studies have been '*ex post*' or 'retrospective', in other words assessing completed research.

Expenditure on research is an investment, expected to generate returns that are sometimes called the payback from research. Traditional economic methods for assessing returns on investment, based on market prices, are inadequate in the context of research. Research conducted in or funded by the public and not for profit sectors, by definition cannot be valued simply by its effect on commercial activities or its market price, not least because the outputs of publicly funded research may be available freely as a public goods.

Even in the private sector, special tools may be required to analyse returns from this type of investment since the relationship between investment in research and its payback may be complex and unexpected, and typically occurs over an extended time period. Moreover, the returns to individual companies or even industries are likely to underestimate the benefits of their research, if commercial benefits spill over between sectors, and if the research generates non-commercial benefits.

This chapter outlines the range of different types of payback that might be generated by health research. It surveys particular issues that arise when selecting a method to measure payback, and outlines methods commonly used to assess payback: review by experts or users; bibliometrics; economic appraisal; and social science methods. These methods are not mutually exclusive, but rather are presented as a range of alternative approaches.

Categories of research payback

Research produces value, or payback, across a wide spectrum from creating knowledge to saving lives to stimulating new industries and to providing employment. Earlier work has identified a range of benefits from research (see, for example Gibbons and Georghiou, 1987) and specific analysis of health research validated a categorisation of the different dimensions of payback (Buxton and Hanney ,1996).

There are five broad categories of payback, summarised in Table 1. The first category, Contribution to Knowledge, is valuable both insofar as the production of knowledge is an end in itself, and if the knowledge is likely to contribute in the future to the production of other types of payback.

The second category comprises benefits to research processes and the organisations conducting research. These paybacks accrue if research refines research methods and competencies, either technical or methodological. They may also accrue if the results of one research project refine the parameters for future research, and if the skills, networks and knowledge acquired by researchers increase as they conduct research. Furthermore, there is some evidence that organisations engaged in research are more likely to use outside research findings and exploit the results of scientific breakthroughs (Gambardella, 1992). Some commentators believe that being involved in research is an important part of staff development and may increase the ability of NHS staff to adapt their practice in line with research findings (Elderidge and South, 1998).

The third category of payback constitute political and administrative benefits. There may be benefits if research contributes to an improved information base for decisions which can range from those taken by national policy makers to those from local administrators and royal colleges. There are sometimes other benefits to politicians, public servants and other decision-makers. These include commissioning research in order to be seen to be 'doing something'; in order to delay decisions; or to legitimate decisions actually taken for other reasons.

The fourth type of payback involves health gains. Research may bring direct returns if it reduces the cost of providing health care, allowing health gain by releasing resources that can be used to better effect elsewhere. Research which leads to intellectual

Table 1 Categories of Research Payback

Payback category
a. Contribution to knowledge
<p>b. Benefits to future research:</p> <ul style="list-style-type: none"> i. future research can be more precisely defined and targeted, in either method or scope; ii. increase the skills, knowledge and professional networks of researchers; iii. enhanced ability to utilise or capture existing research. iv. staff development/educational benefits
<p>c. Political and administrative benefits:</p> <ul style="list-style-type: none"> i. improved information base for decision-making, which is assumed to be beneficial even if particular research findings are not ultimately implemented; ii. benefits to decision-makers including being seen to take action, delay decisions, or justify decisions taken for other reasons.
<p>d. Health benefits:</p> <ul style="list-style-type: none"> i. reduction in the cost of delivering services; ii. improvements in the quality of service delivery, including increased patient and staff satisfaction; iii. improved efficacy of health services, leading to increased health gain; iv. improved equity, by improved allocation of resources across areas, better targeting and improved accessibility of services. v. revenues to NHS from Intellectual Property Rights
<p>e. Broader economic benefits:</p> <ul style="list-style-type: none"> i. commercial exploitation of innovations; ii. productivity gains resulting from a healthier workforce.

Source: Buxton and Hanney 1997.

property may generate income for the NHS, although it should be noted that this constitutes a transfer rather than an absolute gain. Research may also bring qualitative improvements to the process of health care delivery, such as reductions in waiting times, in turn improving patient satisfaction with services and/or staff satisfaction. Research may improve the effectiveness and efficacy of health care, thus resulting in a health gain. It is also feasible that some types of research might improve equity of access to health services, by improving the ability to allocate financial resources, by allowing improved access to health services, and by addressing problems particularly relevant to the least well-off in society.

The fifth type of payback comprises economic benefits accruing outside the health sector. R&D which leads to the development of new products or processes may have a direct impact on economic growth and on employment. Non-commercial health-related research may also have an indirect impact on economic indicators, if it improves productivity by, for example, reducing sickness-related absenteeism.

General issues in the measurement of payback

A number of issues must be considered when selecting a measure of any or all of these dimensions of payback. Three are discussed briefly in this section. Later chapters provide detailed discussion of how these issues are manifest when measuring Department of Health R&D.

Measurement of outputs versus outcomes and the development of a payback model

Traditionally a distinction has been made between output indicators which measure the direct products of research, and outcome indicators which measure changes in variables directly related to health and economic benefits. Examples of output indicators include the number of papers published and disseminated at conferences, and the number of completed PhDs. Outcomes such as health benefits are usually harder to measure directly. Therefore, output indicators might usefully be used as accessible indicators of outcomes *if* there is evidence of correlation between the two.

It can be helpful to model the process whereby R&D generates final outcomes. Buxton and Hanney (1996, 1997) use a sequential model, starting with the identification of R&D needs. There is then the first of two interfaces between the political and professional environment and the research processes. The permeability of the interfaces can be important if outcomes are to be achieved and the first interface involves the commissioning and specification of research to meet the identified need. The primary, or initial, outputs emerge from the subsequent research processes and include publications. There is then the second, or dissemination, interface when findings are relayed to the political or professional environment. This is followed by a stage which may include the generation of secondary outputs in the form of policy or administrative decisions, manifest in, for example, national ministerial decisions or local guidelines. Final outcomes often result from the application of these secondary outputs, and comprise health and economic benefits. Although in its simplest form this is a linear model, it can incorporate and highlight the variety of feedbacks

that occur, and can be contextualised within a model recognising the existence of a 'knowledge pool'.

Quantitative versus qualitative assessment methods

Indicators of both the output and outcomes of research may be quantitative or qualitative. Contribution to knowledge may, for example, be measured quantitatively by the number of papers published, or qualitatively by expert review.

The choice of qualitative or quantitative method depends on the available indicators and on whether relevant dimensions can be captured by available quantitative indicators. In practice, qualitative methods are often more costly to implement, and may generate results that are less comparable across different projects or institutions. The lack of comparability may, however, result from inherent differences between projects, differences that would be masked by a quantitative indicator.

Attribution, time lags and the 'contributive role'

It is often not straightforward to attribute a given outcome or even output to its component research-inputs or, conversely, to track the outcomes and outputs of a given piece of research. Isolating the payback from one research project, or even one research institute, is difficult since research is often a collaborative, cumulative process. Any one breakthrough in, for example, medical practice, is likely to draw on many different research projects. Conversely, one research project may have multiple impacts. The key question is: would a particular payback have been achieved without a particular piece of R&D?

Attribution is also complicated by the considerable time that may pass before a research project bears fruit. The now classic study, Project Hindsight, which examined the relationship between scientific research and innovations in defence, found that up to 20 years may elapse between the publication of research results and realisation of applications (Sherwin and Isenson, 1966).

The time lags mean that full assessment of payback will take place either after considerable time has elapsed or under conditions of uncertainty about the eventual level of impact. In other words, interim assessment of payback will have to be based on the likelihood (or probability) of particular outputs and outcomes.

Methods of measuring payback

The issues outlined above arise regardless of the method used to measure payback. A number of methods are currently employed, in various contexts. Four of the most commonly used types of method are described briefly in this section. Table 2 summarises the methods, and their advantages and disadvantages.

Expert review, by peers or users

Expert review may be used to judge the value of individual publications, or of whole research programmes. Peer review is a traditional means by which the value (especially in terms of scientific quality) of research is assessed. It is undertaken by one or several experts in a field, able to judge the contribution to knowledge made by particular research or research groups. It provides a means whereby expert technical judgement can be employed. This is particularly useful when outcomes are intangible or multidimensional, or when judgement is required to identify the particular contribution made by one research project to a complex outcome.

Review may also be undertaken by the users of research, defined as those who might, by virtue of their own responsibilities and interests, be expected to use the results of the research. They may include health care policy makers, health care professionals, industrial representatives, and members of patient groups. A variety of different techniques may be used to gain the views of users, including panel discussions, interviews, and questionnaires.

Table 2 : Characteristics of different approaches to assessing payback*

Source of information	Research type	Strengths	Weaknesses	Relative cost
Review - expert/technical	All types	<ul style="list-style-type: none"> • can provide information on a range of dimensions of value and of the role of different researchers and organisations. 	<ul style="list-style-type: none"> • may be a burden on reviewers; • may be biased, because relies on information from a small number of individuals; • only qualitative information or, if use scoring methods, these may be hard to validate and not transferable. 	low total cost, high cost to individual reviewers.
Review - by users	Applied and developmental	<ul style="list-style-type: none"> • can provide comprehensive information about a range of valued dimensions; • if use surveys, overcome the small numbers problem. 	<ul style="list-style-type: none"> • considerable information may have to be given to reviewers; • qualitative information or non-transferrable scoring system; • may be hard to develop and analyse survey - results may be ambiguous; • users may be unable to attribute results to projects or programmes. 	medium (development and analysis of surveys may be difficult).
Questionnaires to researchers	All	<ul style="list-style-type: none"> • can provide comprehensive information about value. 	<ul style="list-style-type: none"> • may have low compliance, since places an administrative burden on researchers; • responses may not be accurate, may need verification; • researchers may not have full information about all uses of and outcomes from their research. 	Medium, depending on level of validation required (does impose costs on researchers).
Bibliometrics (publications)	Basic, applied (some)	<ul style="list-style-type: none"> • information is readily accessible; 	<ul style="list-style-type: none"> • only a partial indicator of payback; • may be biased by confounding effects of differences between subject areas and types of research. 	medium-high (depending on type of analysis conducted).
Economic assessment	Applied and developmental	<ul style="list-style-type: none"> • structured and comprehensive framework. 	<ul style="list-style-type: none"> • considerable data requirements; • results dependent on assumptions, especially given the level of uncertainty and extended period over which benefits accrue; • many benefits not able to be quantified. If use cost-effectiveness analysis, can't compare across programmes; • only feasible to apply to a subset of R&D projects. 	high (data collection and analysis)
Case studies	All	<ul style="list-style-type: none"> • provides information about causality, and about the relationship between R&D and its impacts; • provides comprehensive information about all dimensions of payback. 	<ul style="list-style-type: none"> • results may not be generalisable; • hard to compare across different case studies. 	Medium (vary with number of case studies).

* Based on Williams and Rank, 1998.

A number of different methods can be employed in the review process. It may involve examination of publications, site visits, and interviews. The process may be informal or structured. If it is structured, reviewers may be requested to give particular qualitative information or to make quantified ratings (Cozzens, 1997).

Bibliometric techniques

Bibliometric analysis uses data about publications or patents to analyse the payback from research. Bibliometric analysis of payback commonly comprises comparison of groups of researchers in terms of their activity, measured by the number of publications and type of journal in which publications appear; and their impact, measured by the number or type of citations. (Narin, Olvastro and Stevens, 1994). An early, classic study measuring the linkage between articles and citations, assessed the influence of basic scientific research on medical breakthroughs by tracing citations. (Comroe and Dripps, 1974, 1976).

There are a number of public datasets suitable for bibliometric analysis, including the Science Citation Index (SCI), the Social Science Citation Index (SSCI), Medline and so on. These datasets comprise information about all papers appearing in a very large number of journals over a period of years. The information held in the datasets includes information about the authors (including their addresses); keywords; information about the journal; and information about the number of citations of each paper and where they appear. The Wellcome Policy Unit has a database, the Research Outputs Database (ROD), which includes all of the information contained in the SCI, as well as information about funders where this is included in the publication. (Lewison, 1998).

Analysis sometimes uses a Journal Impact Factor (JIF), which is a score calculated for each journal, based on the total number of citations of all papers appearing in that journal during a particular period, weighted by the number of articles it publishes. By this measure papers published in journals with high JIFs (for example, *Science* and *Nature*) are deemed to be more valued than papers in less frequently cited journals.

Bibliometric analysis using published datasets has the advantage of imposing few direct costs on researchers. It can therefore be applied across a large number of groups of researchers to compare the output of, for example, all of the hospitals in the UK. If used routinely it may create a particular incentive structure. The use of JIFs, for example, rewards research with a greater likelihood of being published in journals with a high JIF. This may be a desirable or an undesirable consequence, depending on the objectives of decision makers.

There are also a number of disadvantages to bibliometric analysis (Adams, 1997; Herbertz, 1995). There are practical difficulties in using the databases: they are not 100% accurate or complete; and they cover only a subset of journals, with relatively poor coverage of, for example, health services research journals (Black and Davies, 1999) and journals that are not published in English. The normal number of citations and publications varies between subject areas, meaning that direct comparisons can be made only after controlling for differences. Moreover, publications are only one output of research, and they are an output not useful in themselves but useful only insofar as they indicate a contribution to knowledge, improvement in research capacity, or expansion in the information base available to decision-makers.

Economic assessment

This section provides a very brief survey of a number of the different methods of economic analysis that can be used to measure a number of the dimensions of payback from research.

Econometric studies use multi-variate statistical methods to estimate the contribution of research expenditure to economic growth (at a national level) or profit (at the level of individual industries or companies). In these studies, the value of research is taken to be its contribution to economic growth, productivity or profit.

Various investment appraisal techniques measure the rate of return on research as an investment, by comparing a quantitative measure of benefit against cost. For example, one study of the return to private sector research found an average internal rate of return of 28% (Mansfield, 1991). As with economic evaluation, these techniques usually discount costs and benefits which, given the extended period before benefits accrue, leads to results with apparently low returns.

Portfolio analysis combines a number of different projects by using programming techniques to find an optimal allocation of resources between the projects. By allowing a number of risky projects to be combined, portfolio analysis provides a means of managing uncertainty.

Option appraisal treats investment in research as the purchase of an 'option', opening future possibilities. In other words, developing an R&D capacity gives the developer the option of future involvement in particular areas.

Economic evaluation allows systematic comparison of costs and benefits. Benefits may be measured by a particular health gain, such as a change in life expectancy, by QALYs, or by monetary impact. Weisbrod, for example, conducted retrospective analysis of the value generated by the research leading to the successful development of a vaccine for polio in 1957. He estimated that by 1971 the total monetary value of the treatment costs saved and the reduction in mortality and morbidity was \$US 1,350 per case prevented (1957 prices), equivalent to a return on research investment of 11-12% (Weisbrod, 1971). In practice, economic evaluation is difficult to undertake, given the cost and problems of acquiring relevant information and estimating benefits.

Social science methods

A number of methods are used by social scientists to estimate payback, including case studies and questionnaires.

Case studies can be defined as 'the systematic documentation of the impacts of R&D that has been carried out in the past' (Williams and Rank, 1998, p53). Case studies may be undertaken as an end in themselves, or may be undertaken to provide information to be used as an input to economic appraisal. They have the advantage of being able to provide information on a range of different dimensions of payback, potentially on all of the dimensions outlined above including the impact of R&D on policy, and can be used to establish causality. Case studies are, however, resource intensive and so are unlikely to be appropriate as the main method in a routine monitoring system (Buxton *et al*, 1999).

Questionnaires are commonly used to gather information on 'payback' from the researchers themselves, and even where formal questionnaires have not been used, many funders have adopted some system of structured final reports which may act, rather like a questionnaire, as a source of information on initial publications. Considerable effort may be required to validate the instrument and to analyse results.

Nonetheless questionnaires are still likely to be less resource intensive than case studies, making them more suitable for routine monitoring. Response rate may be poor if it is voluntary, given the cost in researcher time required to fill in questionnaires. If researchers think there is a link between their answer and their funding, response rates may be higher, but there may then be a danger of bias.

Questionnaires can also be used to gather information from users of research and can be administered to a large number of people routinely. The Economic and Social Research

Council (ESRC) is undertaking a major study of the use of user panels and network analysis, to assess the impact of research on non-academic users. Exploratory work showed that users are likely to have difficulty in attributing research findings to particular bodies or projects, suggesting that users may not be able to assess all types of R&D (Cave and Hanney, 1996; ESRC, 1997).

Conclusions

The choice of a method for measuring payback depends, ultimately, on the objectives of the evaluator, since different methods highlight different dimensions of payback. Clear understanding of policy objectives, in the context of which categories of payback are most important, must therefore inform the selection of a method. For example, a funder concerned primarily with increasing research capacity and facilitating health gains should not rely solely on bibliometric methods to measure value. As will be discussed in subsequent chapters, the dimensions of value important to the Department of Health can be measured adequately only if a multiple measures are used. This principle underlies the system we sketch in Chapter 6.

This chapter has focused on research leading to the development of methods for valuing payback. A routine monitoring system needs to reflect this research experience, but if it is to be routinely used it must also address particular operational questions, which are discussed in Chapter 3.

CHAPTER 3 : CHARACTERISTICS OF A ROUTINE MONITORING SYSTEM

Introduction

Chapter 2 illustrated the range of approaches that have been adopted in *ad hoc* research studies of payback. As we ~~move examine the possibilities of moving~~ towards a system that might routinely monitor the payback from a funding agency's health research, ~~we need to focus on three dimensions. the thinking needs to push forward in a number of directions.~~ Firstly, ~~we~~ needs to move towards an approach that can be ~~routinely~~ undertaken routinely. Secondly, ~~we~~ needs to accommodate focus on outcomes as well as outputs ~~and outcomes~~. Thirdly, ~~it-we~~ needs to find measures which can encompass the full list of payback categories.

In considering ~~the what are~~ appropriate characteristics of a monitoring system it is useful to begin by examining the wide range of assessment approaches that are applied to public services in general. A range of bodies have been active in this field, for example, in the UK, the Audit Commission, the National Audit Office (NAO), and the Treasury and ~~in from~~ the USA, the Office of Management and Budget. Some previous and ongoing attempts to apply monitoring systems to research also offer useful insights lessons, particularly into the problem of attribution when analysing basic research. Overall, the chapter explores the general criteria and necessary characteristics ~~offer~~ a monitoring system.

Assessment approaches and characteristics of systems

Peer review has long been an integral part of the process of assessing research proposals and products. However, ~~the this~~ tradition of evaluating ~~ing ion of~~ research now has to be put into the context of a general growth of interest in output and performance measurement in the public sector in the UK (~~see, for example Cave et al, 1990~~) and elsewhere (see, for example, Cave et al, 1990; Treasury Select Committee, 1999). This expansion of performance measurement and indicators is just one facet of a general development in many Western democracies of what has been referred to as 'New Public Management' (~~see, for example~~ Pollitt, 1993). It has resulted in a situation where there is now a wide range of overlapping approaches to assessment in the public services which include: evaluation; audit; output and performance measurement; and performance indicators (PIs). ~~One example of overlap is that (PIs performance indicators~~ can be incorporated into, or used by, any of the other approaches.) ~~Furthermore, there are debates about how far traditional programme evaluations are converging with performance measurement or audit (see, for example, Keen (forthcoming); Wisler, 1996, and, specifically in relation to research, Cozzens, 1997).~~

The term 'evaluation' covers a wide range of activities but generally ~~it~~ involves making subjective judgements based ~~on data~~ to varying degrees on data. In relation to research, evaluation ~~it~~ is often based on peer review, and whilst such evaluations are sometimes routine, they rarely examine outcomes. The term 'audit' also covers various activities including financial audits, which concentrate on issues such as probity, and the NAO's Value for Money audits, where it is sometimes claimed that the ~~three~~³ E's of economy, efficiency and effectiveness are important (though Keen, forthcoming, disputes this). Such audits can involve checking that planned activities have been completed. 'Medical audit' is different again: in an exchange in the Public Accounts Committee Alan Langlands agreed that the term medical or clinical audit involved either self-assessment or peer review as opposed to the third party activity normally associated with audits (HC 304, 1995-6).

The Value for Money reports from the NAO or the Audit Commission generally involve the production of many blocks of information or PIs, upon which judgements are made. The Audit Commission conducts follow-up studies which track progress using indicators based on the original study recommendations. A recent example involved the Accident and Emergency Services (Audit Commission, 1998), where independent judgements from the auditors, such as 'there is still scope for better deployment of staff' (p1), are made on the basis of the indicators.

As a result of government pressure there was a large increase in the number of output and performance measures listed by departments in the Public Expenditure White Papers - rising from 500 in 1985 to about 2,300 in 1990 (Treasury, 1990). Further work by central government on these issues throughout the 1990s has seen several strands of thinking including the use of performance targets. The development of resource accounting has led to the situation where in 1999-2000 UK government departments produce 'their first set of Output and Performance Analyses' (McConaghy, 1999, p62). There has also been the development of Public Service Agreements (PSAs) which include a series of performance targets for each department which are 'wherever possible, "SMART" - specific, measurable, achievable, relevant and timed' (Chief Secretary to the Treasury, 1998, p5). The Treasury Select Committee reported an estimate 'that two thirds of the PSA indicators were of outcomes' (Treasury Select Committee, 1999, para 14). It also reported a Treasury assessment that there were about 600 PSA targets but about 1000 Output and Performance Analyses targets. The committee argued that although the two had a separate history it would be worthwhile amalgamating 'the two documents on the next occasion on which a full set of PSAs is published' (para 22).

It is certainly desirable to monitor a system against its objectives. However, as discussed below, it is unlikely that a monitoring or an audit system for research output and outcomes will be able to be assessed against precisely defined quantifiable targets.

It might be reasonable to conclude from this section that, at this stage in our understanding of payback from research, we should be looking to develop a system of outcomes monitoring that incorporates performance indicators and measurement, rather than an audit system that is ~~trying to~~ ~~that could~~ monitor activities against predetermined quantitative targets.

General characteristics of performance measurement systems and their relevance to research

In the development of performance and output measurement systems there are various indications that research is regarded as a rather separate area of public expenditure. When describing the use of output and performance measures in government departments Lewis and Jones (1990) recognised that different types of PIs were appropriate for the different types of activity performed in government departments. But in listing specific PIs for R&D establishments they noted that these PIs were 'mostly directed towards measuring the performance of R&D *administration*' (p51). The difficulty of including research outputs in performance audits or measurement systems was again highlighted by the NAO report into performance review in the Defence Evaluation and Research Agency. It noted that Treasury guidance to agencies recommended targets on financial performance, volume of output, quality of service and efficiency. The NAO report continues, however, without demurring, 'The guidance singles out research establishments as being examples of agencies with non-standard outputs: it makes little sense to count the numbers of reports issued where a 'report' is not of constant size, cost or value' (NAO, 1997, p71). The NAO report, nevertheless, did undertake some analysis to show the annual production of scientific papers per research leader. They also examined the Agency's customer satisfaction surveys and undertook some customer focus groups of their own.

A review of research evaluation in the UK noted that research is funded by different government departments for differing purposes and, therefore, claimed that different types of evaluation would be appropriate. Of the Department of Health's research it noted that 'this relatively complex system militates against a uniform approach to evaluation' (Hills and Dale, 1995, p40).

Difficulties in dealing with research in a general performance monitoring system have occurred in the USA where the 1993 Government Performance and Results Act (GPRA)

required federal agencies to set strategic goals and to use performance measures for management and budgeting. The Committee on Science Engineering and Public Policy (COSEPUP) recently reported that 'the specific goal of GPRA is to focus agency and oversight attention on the outcomes of government activities [but] Development of plans to implement the act has been particularly difficult for agencies responsible for research activities supported by the federal government because of the difficulty of linking results with annual investments in research' (COSEPUP, 1999). ~~According to Cozzens (1997) many research agencies hoped that they would be exempt from the GPRA, but in fact this is not going to be the case. Nevertheless she explains why 'research outcome measures are not, and in fact will not, become available for GPRA-type performance reporting' (1997, p78).~~

Following an analysis of the differences between traditional evaluation of a research programme and the requirements of the GPRA, Cozzens (1997) lists various indicators that would 'satisfy the standard GPRA template' (p86). They are: publication counts; citations per publication; doctorates produced; and user involvement and satisfaction ratings. She then comments that, 'The problem with the set, of course, is that it leaves out virtually all of what researchers themselves find important about their work' (p86). As a way out of the difficulties she suggests that: 'many agencies are now turning to the alternative goal-setting and reporting format available under GPRA. This alternative format allows descriptive performance goals, requiring only that agencies paint a word picture of the difference between minimally effective and successful performance. The match between actual performance and the word picture is subject to the same requirements for evidence that apply to quantitative goals. This path thus provides both the flexibility research agencies need to maintain a focus on outcome rather than outputs of their process, and the rigorous standards GPRA encourages' (p88).

This discussion highlights the difficulties in trying to develop a set of research output and outcome indicators that fit into standard approaches to output and performance measurement. This suggests that a highly specific monitoring system might have to be developed for the outcomes of health research. Furthermore, it is likely that such a system might be relevant for only some types of research. Cozzens' analysis seeks to show the advantages in the research context of traditional programme evaluation or assessment over summary PIs. She describes how research outputs enter a knowledge pool (Gibbons and Johnston, 1974) in which ideas and people interact and produce innovation through unpredictable paths which makes the tracking of eventual outcomes particularly difficult to do other than through expensive retrospective studies. The analysis of Buxton and Hanney

(1997) and Hanney *et al* (forthcoming) acknowledges the complexities caused by the stock or pool of knowledge but also describes more readily identifiable flows of knowledge that can sometimes occur, especially with policy research or technology assessment.

For basic research programmes, COSEPUP argues that it is necessary to evaluate programmes 'by using measures not of practical outcomes but of performance, such as the generation of new knowledge, the quality of research, the attainment of leadership in the field, and the development of human resources'. These could best be evaluated, it recommends, by expert review and by using the alternative descriptive GPRA format outlined above. Both the National Science Foundation (NSF), and the National Institutes of Health (NIH), however, make a distinction between outcome performance, to which the above analysis applies, and processes or means where more quantitative approaches against targets are recommended. The NIH has been allowed to have its performance assessed at a higher level of aggregation than the NSF.

For applied research COSEPUP suggests that the requirements of the GPRA could be met by annually measuring progress towards practical outcomes. It recommends setting outcome targets in the form of milestones. For example, a programme might be set the milestone of adapting fibre-optic laser surgery for the treatment of prostatic cancer, or of devising a prototype DNA sequencing machine (COSEPUP, 1999). There are, however, likely to be significant problems when trying to set appropriate milestones for health R&D programmes. Although much DH/NHS R&D is applied, it is unlikely to be the case that precise, verifiable outcomes can be specified for programmes which cover a wide range of topics. In practice, it is easier to set milestones that relate to process rather than to outcome variables.

Some commentators argue that it is impossible to measure the value of R&D. There is, however, convincing evidence that it is feasible to measure value, for at least some types of health-related R&D ([Buxton and Hanney, 1997](#)). It is not, however, clear that the same indicators can be used to measure the value generated by all types of R&D. As will be discussed in Chapter 4, it may not be possible to use the same set of indicators to measure the value of basic bio-medical science, health services research and social care research. This means that, in the context of DH funded R&D, a comprehensive routine monitoring system must comprise a variety of different indicators and assessment techniques. A system of this type might be called an Outcomes Monitoring System. As will be discussed in Chapter 6, which outlines the features of such a system, combining different indicators presents substantive analytic problems. Not combining them would, however, sacrifice

comprehensiveness for tractability. The next and final section of this chapter describes the criteria against which a routine monitoring system must be judged.

Criteria for a routine monitoring system for R&D

Various criteria have been developed for judging the adequacy of individual performance indicators for higher education (Cave *et al*, 1997; Sizer, 1979). These criteria can be adapted so they are suitable for judging an outcomes monitoring system for health-related R&D. A system will provide information useful for managing R&D if it meets the following criteria:

- * *relevance to objectives*, in that it measures the full range of dimensions relevant to the funders' objectives. It should be as comprehensive as possible;
- * *decision-relevance*, in the sense that the information can be analysed and presented in a way that is consistent with the funders' decision making processes;
- * *truthful compliance*, in that the system should encourage those holding information to comply with requests to reveal the information at all, and to reveal it accurately. This can be achieved either by giving appropriate incentives to those holding information, or by monitoring their returns to check their completeness and accuracy;
- * *minimises unintended consequences*, in other words has incentives rewarding desirable behaviour. The previous criterion, truthful compliance, refers to behaviour related to giving information. This criterion refers to behaviour related to the R&D activities themselves;
- * *acceptable costs* relative to its benefits.

The routine evaluation of research in universities in the UK conducted through the Research Assessment Exercise (RAE) powerfully illustrates the importance of considering incentives. Although the RAE assessment process has evolved over time (~~see, for example, Cave *et al*, 1997~~) in every case it has involved peer review of departments informed by review of publications and data on research students and on income. Its importance in terms of research resources to departments is so great that it intensifies the pressure on academics to give maximum attention to the quality of their research as measured by RAE, and as a result to be less concerned with other research outcomes—resulting from the research findings. It is therefore important that any system for evaluating NHS research should

recognise its interactions with the RAE. ~~Secondly, the specific indicators used in the RAE provide a good example in the discussion below of how incentives can operate.~~ NHS and university R&D are not always distinct. 60% of NHS publications involve a collaboration with at least one university author.

The specific indicators used in the RAE provide a good example of the operation of incentives. Performance indicators form a vital part of the incentive structure, insofar as they connect behaviour with rewards or penalties. Incentives operate by rewarding or punishing particular courses of action. They may be financial or non-financial, since people have objectives relating to both of these types of reward (or penalty). Non-financial incentives may relate to private self interest, for example, a desire for leisure time or peer approval. They may also related to public interest, for example the desire of public officials and medical professionals to fulfil their agency role to patients or citizens (Scott, 1997; Wilson, 1989).

Performance measures may be designed to be part of incentive systems and have intended consequences, such as the payments to GPs for health promotion activities, designed to stimulate increased activity. They may also have unintended consequences. In the context of R&D, incentives similarly reward particular behaviour, in both intended and unintended ways. For example, for the 1992 HEFCs' Research Assessment Exercise the impression was gained by researchers that the number of publications would be used as a performance indicator. This led to publication inflation (O'Brien, 1994).

More generally, unintended consequences may result from using PIs which only measure the easily measurable dimensions of research outcomes. Focusing on these

~~measurable dimensions gives individuals an incentive to focus on activity affecting these dimensions, which may mean that they ignore other, equally important dimensions. More generally the fact that only a subset of the dimensions that are important can be observed and easily measured means that it may not be optimal to use as performance indicators those that could be applied. To do so might lead to undesirable behaviour, as individuals respond to the indicators, ignore other dimensions that are as important.~~

Conclusions

The issues outlined in the first parts of this chapter mean that the value of health-related R&D cannot be measured adequately using a conventional set of PIs. As will be discussed further in Chapter 4, a system based on a single set of conventional PIs is unlikely to be able to accommodate all types of R&D, from basic biomedical science to health services research. Conventional performance indicators are not, therefore, sufficiently comprehensive to capture key dimensions valued by funders such as the DH. A more complex monitoring system will be necessary.

Appropriate approaches therefore need to be developed for application to health R&D. Chapter 6 sketches an outline of an outcomes monitoring system that potentially meets the criteria outlined in the latter part of this chapter: relevance to funders' objectives; decision-relevant; accurate; minimises unintended consequences; and with acceptable net costs.

CHAPTER 4 : DIFFERENCES BETWEEN RESEARCH TYPES

Research and development activities may be categorised in a number of different ways. Differences occur on a variety of dimensions: sources of finance; types of knowledge generated; norms of information disclosure; reward systems (Dasgupta and David, 1994). Recognising the differences is important to the design of the routine monitoring system sketched in Chapter 6, since different types of research may perform differently on any particular indicator. This chapter discusses first, differences between basic, applied and experimental R&D; second, the special case of social care research; and third, differences between commissioned and responsive research.

'Basic' versus 'applied' research

The OECD distinguishes three types of R&D activity (OECD, 1994):

- basic research is defined as 'experimental or theoretical work undertaken primarily to acquire new knowledge of the underlying foundations of phenomena and observable facts, without any particular application or use in view';
- applied research is defined as 'original investigation undertaken in order to acquire new knowledge. It is, however, directed primarily towards a specific practical aim or objective';
- experimental development is defined as 'systematic work, drawing on existing knowledge gained from research and practical experience, that is directed to producing new materials, products and devices; to installing new processes, systems and services; or to improving substantially those already produced or installed'.

In the context of health-related research, for example, basic research includes mapping the Genome; applied research includes relating particular clinical symptoms to particular genetic patterns; and experimental development includes developing a routine test or a treatment for the clinical symptoms.

Using the OECD's definitions, the Office of Science and Technology (OST) estimate that in the UK in 1996/7, basic research accounted for 32.1% of total government R&D expenditure; applied research 40.6% and experimental development 27.3% (OST, 1998). In the same year, basic research accounted for 17% of DH R&D expenditure; 63% of the MRC's expenditure; and none of the NHS's R&D expenditure. An alternative estimate of the balance between applied and basic R&D can be made by examining the type of journal in

which publications funded by public organisations appear. Work currently underway at the Wellcome Trust Policy Unit suggests that in 1996 at least 54% of NHS-funded work was published in applied journals and at least 32% in basic science journals (Personal communication, M.Yare and J.Grant). These estimates suggest that both are important in the DH and NHS, and that both need to be accommodated in any comprehensive routine monitoring system.

Basic, applied and experimental R&D may vary in any or all of the dimensions listed above. Basic science and some types of applied research are more likely to be characterised by knowledge outputs that are placed in the public domain (Dasgupta and David, 1994). Experimental development is, by contrast, sometimes intended to develop tradable products or applications in which case it is patented or kept secret.

Although they differ in important respects, different types of R&D are complementary parts of one process. Basic research can be shown to play a vital role in, for example, breakthroughs in medical treatment and diagnosis (Comroe and Dripps, 1974, 1976). And conversely, the questions faced in experimental and applied work can usefully inform basic research.

The process of information flow between different types of R&D is complicated by and sometimes impeded by their differences, particularly differences between their goals, norms and reward systems, and by barriers between organisations. Current government policy therefore seeks to facilitate collaboration, increasing the rate at which the results of basic research inform applications and facilitating the process by which basic research is informed by applied questions. This is evident both in policy relating to the delivery of health care and in research policy. In the context of research policy, collaboration is encouraged between basic scientists and those able to develop the results (Our Competitive Future, 1998). Collaboration is also fundamental to recent reforms of NHS R&D funding (Supporting R&D in the NHS 1994; Culyer, 1996).

These factors are important to the design of a monitoring system. Indicators form part of the incentive structure, rewarding or penalising particular activity. The choice of indicator will, therefore, affect the likelihood of collaboration occurring. It is possible to design a monitoring system with indicators that directly measure (and therefore reward) collaboration. This is already common among funding bodies, which make *ex ante* evidence of collaboration a prerequisite to some types of funding (Landry and Armara, 1998). Indicators should also be consistent with the outputs of each type of research if they to be consistent with collaboration.

In general, differences between different types of R&D affect what will be measured. The different types of R&D generate different outputs and outcomes, and may therefore perform differently on the same measures. As will be discussed in Chapter 6, recognising this is the case is central to designing a valid monitoring system.

Performance on bibliometric indicators, for example, will be affected by differences in norms about whether, where and what to publish, meaning that variation in publications cannot be interpreted as necessarily indicating variation in payback. Social scientists, for example, tend to have fewer publications, containing larger 'chunks' of research, than natural scientists. Variations in Journal Impact Factors (JIF) between fields may similarly reflect differences in the outlets used by different types of research, or differences in the size of a field. These factors mean that bibliometric indicators of payback can be used to compare different fields or types of research only after confounding factors have been controlled.

Reservations about the use of bibliometric indicators are particularly important in the context of health-related research, since relevant R&D occurs across a number of different subjects and in all three broad types of R&D.

Moreover, important differences between the incentives governing basic as opposed to applied researchers mean that bibliometric indicators should not be used as the sole basis for comparing all types of research (Black, 1997; Dasgupta and David, 1994). Employing publications as an indicator is consistent with desirable behaviour by basic scientists, since this is a measure of whether knowledge is placed in the public domain. Publications are, however, inconsistent with the incentive structure governing many clinicians, since they are often not rewarded for publications. Unless they have a university contract, their employment is generally not conditioned on their R&D activities (Fears and Poste, 1998). Hospital medical consultants, for example, are contracted to undertake patient services, not to undertake R&D (Elderidge and South, 1998; HSMU, 1999).

Whereas publications reward making knowledge public, patents effectively 'privatise' knowledge. They therefore allow innovators to appropriate returns on their investment in R&D, perhaps giving them an incentive to make that investment. Patents may, therefore, promote desirable behaviour in some types of researcher, particular those engaged in experimental development. They are not, however, desirable in all contexts since many types of research do not lead to patentable results, nor is it clear that privatisation of applied knowledge promotes efficiency in health care.

Table 2 summarises the measures of payback, including bibliometric measures, that are suited to different types of research. Expert review is suitable for all types of research. User review is feasible only for applied and experimental R&D since, by definition, the users of basic research are not likely to be known. Questionnaires to researchers are relevant to all but may need to be adapted to the particular type of research in question. Bibliometric measures are particularly relevant for basic research and some types of applied. Economic appraisal is unlikely to be feasible for monitoring basic research, given the uncertainty about the eventual uses of the research and the long time lags involved. Moreover, the data required for meaningful economic appraisal mean that it can be used only for detailed examination of a few studies. Case studies, involving detailed investigation of the outcomes and outputs of particular R&D projects may be useful for all types of R&D.

Social care research

For some health research funding agencies there may be a need to consider applied research that is not concerned with health *per se*, for example research on aspects of education or social care. Do these raise special problems? Taking the example of research on social care, it would appear in principle that they do not. Reviews of social care research have indicated that it has had impact on policy and practice (Kogan and Henkel, 1983; DH, 1994). It has also proved possible to apply the ~~HERG~~-general approach used to value DH research to some social care projects and programmes (Henkel, 1994; Youll, 1997).

Inevitably some of the categories of health benefits and detailed ways of assessing them will be different for social care research than for health. Furthermore, some items suggested in Chapter 6 may not be relevant. The variations in publication patterns discussed above are also relevant to social care research where books can sometimes be the major publications outlet. Nevertheless at this stage it seems likely that the broad approach proposed in Chapter 6 could apply to social care research. *Applied* social care research may have greater similarities with *applied* health research than applied health research has with basic biomedical science-related-to-health.

Commissioned versus responsive R&D

A distinction may also be drawn between commissioned and responsive R&D. The distinction is relevant to the design of a monitoring system and choice of indicators. For responsive R&D, resulting from open applications to a funding body, funding is given for the investigation of topics chosen by researchers. Commissioned research may take the form of a call by the funding body for research ideas on a particular topic, for example the interface

between primary and secondary care or issues relating to childhood asthma. Alternatively, commissioned research may be even more prescriptive, defining the research question and the method.

Unlike responsive research, commissioned research prescribes the type of knowledge that should be generated by the R&D. As pointed out by Thomas (1985), government commissioned research has built into it some of the minimum conditions for outcomes to be generated. The fact that a particular area is selected for commissioned research suggests that the funder has already identified that a contribution to knowledge in this area will be valuable. In other words, that this is an area where reducing uncertainty by increasing information might generate payback (Claxton, 1998). In this case, the funder requires a monitoring instrument that tracks whether the anticipated payback is achieved. The prescriptive nature of commissioned research means that it should, therefore, generate its own monitoring criteria, in terms of the contribution made to the area of knowledge defined by the funder. In the case of responsive research, by contrast, the expected contribution to knowledge is defined by the researchers themselves. Sometimes it is explicitly part of a funding application.

Conclusions

This chapter has described various different types of health R&D. The different types of research generate different types of outcomes and are likely to perform differently on the same measures of payback, for reasons not necessarily related to the value of the research. Funders' objectives often mean that they are interested in having a balanced R&D portfolio, incorporating a range of different types of R&D. A routine monitoring system must therefore be capable of recognising and rewarding a range of different types of R&D. The system outlined in Chapter 6 has the potential to achieve this.

CHAPTER 5 : WHICH UNITS OF RESEARCH?

Introduction: ~~what are we concerned about?~~

In developing a system for monitoring the R&D activity of a funding agency, there is a need to be clear first, about the basic units of R&D on which data will be gathered, and secondly, the way these may be combined to inform judgements about the relative performance of different parts of the agency's portfolio. These two issues are inter-related.

This chapter outlines the basic types of funding: projects, centres and individuals. It also proposes how data can be collected on outcomes for each of these, in the light of previous attempts to do so.

The main streams of research and the levels within them

In describing the levels at which monitoring could be applied, the diverse use of the term 'programme' has to be addressed. The term is often adopted to describe a collection of projects based on a common theme, ~~for example~~eg, mental health, but funded at a range of institutions. In such a situation, there may be an expectation that interaction and synergy at research and dissemination stages should enhance the outcomes and, to differing degrees, resources are often employed to encourage this. (The label programme is sometimes also used to describe the entire collection of projects funded in the responsive mode, but here there is no project interaction or common themes). Another ~~frequent~~main use of the term is to describe 'block' funding to an institution, centre or research unit, ~~or to describe long term funding to an individual,~~ where the broad subject matter is agreed between the funding agency and the recipients, but the latter have freedom ~~over details relating to the use of the funds as to how the money should be used in detail.~~

To provide an overall framework, whilst avoiding confusion over the use of the term programme, we identify three main streams or modes of research funding. One main funding stream is usually for **projects** administratively grouped into programmes, including a responsive mode 'programme'. The second main stream is funding for **institutions/centres**: this stream includes, in the case of DH/NHS, the 'Budget One' funding for hospital trusts but also includes funding to some university based research units or centres. The third main funding stream is to individual **researchers** as training or career development awards. An agency could have a range of such schemes, for example for PhD fellows, for post-doctoral fellows etc. These three streams are shown in Figure 1 and it is probable that the regular data-gathering for a monitoring system would operate at the basic level of each stream. This is how the analysis in the next section is organised.

Applying indicators to the three streams

It is helpful to consider separately the three streams of R&D activity identified above, and how the indicators discussed in Chapter 2 can be applied at the various levels within them.

Stream 1: Projects in themed programmes or responsive mode

The individual project is the basic level for funding in the responsive mode and for many programmes. Before considering how this level can be viewed as the most appropriate data collection level, it has to be noted that this approach has been subject to criticism on the grounds that it commits the 'project fallacy'. This fallacy, it is claimed, is to assume that individual projects make substantivemuch impact as opposed to possibly playing a contributive role. Nevertheless, most previous analysis of of thethe payback from Department of Health research analysis conducted by HERG has been at the level of individual projects, suggesting that payback categories and indicators are applicable at this level—especially in the two phases of work for the Department of Health (Buxton and Hanney, 1994, and 1997; Ferguson et al, 1998). Most of the payback categories and indicators have therefore been shown to be applicable at this level in at least some cases. In addition, HERG's work for North Thames, and that conducted in other regions has also concentrated on project funding (see, for example, Ferguson et al, 1998). Similarly, a systematic assessments of outcomes of the Quebec Health Technology Assessment (HTA) programme was undertaken at the project level (Jacob and McGregor, 1997). This assessment is of key importance because it found evidence that 18 out of the 21 HTAs it assessed had an impacted on policy. This is a much higher proportion than other researchers usually claim for policy research, and indicates both that HTA might be a particularly fruitful area for examining outcomes and that assessments can operate successfully at the project level.

Previous workHERG's work for the North Thames regional office of the NHS Executive showed that where a programme consists of an organised series of projects on one theme, for example mental health, that are conducted in a range of institutions, the data collection could be at project level, although it could be presented in an aggregate form at programme level. This work showed that aAny extra benefits that may arise, in terms of quality of research and its impact, as a result of projects being part of programmes, should be capable of being gathered in the project level data collection.

But whilst much of the data collection is likely to take place at a project level, expert/user review can be used at programme level. The results of ESRC's examination of the role of user panels for assessing the impacts of its programmes on users could be important here. It is possible however, as shown in HERG's analysis of North Thames Mental Health Programme, for the projects to have impacts but the programme not to be well known to some of the most obvious potential users.

Previous experience of surveying users of Department of Health funded projects has had only ~~had~~ limited success. Although some studies have involved asking departmental personnel about their use of research (~~see, for example,~~ Gordon and Meadows, 1981) it seemed to be at a more general level than in relation to specific projects. Richardson *et al* (1990) suggested the most obvious of the limited strategies available was ~~to~~ survey users, researchers or research managers. Their study advocated developing the pre-existing *ex post* assessment forms which asked ~~the~~ customer divisions about possible benefits accruing from research and whether the research has been used in any of the following categories: legislative changes; administrative changes in advice or guidance; practice action; confirming existing policy or practice action; further research; or active dissemination. They advised the DH on the need for considerable methodological developments and the incorporation of qualitative approaches in any adoption of surveys of use. We understand, however, that these assessment forms are no longer used by the department (Buxton and Hanney, 1994).

~~One~~A problem with adopting user surveys at the project level is that users, even if they are aware of particular findings, might not link them to specific projects from a funding agency. One way round this is to provide users with a list of specific projects and ask them if they know about, and have used the findings from, the specific projects. This approach has been adopted in previous case studies (Buxton *et al*, 1999) with (Bryan, 1997) testing alternative approaches. The dangers of asking users about named projects was highlighted by Drummond *et al* (1996) who reported that some practitioners claimed to have read, and even been influenced by, fictitious economic evaluations included in a survey.

Stream 2: Institutions/centres

The institutional stream is likely to involve the greatest difficulty in relation to multiple funding both from different streams of any specific agency and between different agencies. Funding from one agency at the institutional level can be used in various ways including: for general research infrastructure; for individual project level work to be conducted; to act as core funding that can be used in a leverage way to attract further funding; and for enabling

activities to be funded that increase the likelihood of payback accruing. Questions arise as to how data should be gathered for an institution. Should it ~~include~~ only those activities directly funded by the relevant agency or should it ~~encompass other activities~~~~go more widely~~ when the funding is explicitly seen as having a leverage/enabling role? Overlapping with this is the issue of whether the assessment of benefits from funding in an institution/centre stream should operate at project level or solely at the institutional level. When these two issues (ie. all research at the institution v only research funded by the agency's institutional funding stream; and project level v institutional level) are combined a number of options arise:

- 1) The institution would report at an aggregate level (or at least in a form not related to specific projects) on the paybacks from all the research undertaken at the institution whatever the funding source. This would be similar to some existing practices but would include outcomes as well as outputs. Information would probably be gathered from individual researchers (who would be asked to report on the relevant outputs/outcomes from their work) and not on a project by project basis. There is a danger of double-counting both within, and between, institutions where projects are undertaken on a collaborative basis but where the projects on which the reporting of payback is based are not disclosed.
- 2) Institutions/centres would report on the payback arising specifically from the institutional stream of funding supplied by the funding agency. This could be at an aggregate level or project by project, given that the data would probably be gathered on a project specific basis for those projects using institutional funding.
- 3) Information on payback from any project or individual award being conducted at the centre would be gathered, and reported on, at the project/individual researcher development award holder level. There would clearly be considerable overlap with information being supplied about the other two streams (ie. projects and individual researcher development awards).²

Institutions and Centres increasingly have experience not only of traditional periodic expert review but also of producing annual reports. There has, however, long been disagreement about the ~~dimensions that should be addressed in peer review~~~~most appropriate criteria to include~~, with Henkel and Kogan (1981) describing the debate over whether it should go beyond scientific quality and also examine policy relevance. More recently there have been moves to include a wider range of criteria in reviews of health research centres in the UK (DH, 1992) and in the Netherlands (Spaapen and Sylvian, 1994).

A ~~p~~The previous payback analysis carried out for an NHS region examined two centres. An attempt was made to look at some individual projects or main themes of work as well as at each centre as a whole. However, for reporting purposes it was felt most appropriate to use the aggregate level (i.e. each centre) and here the full range of payback categories was applied (Buxton *et al*, 1999; Hanney *et al*, forthcoming).

Stream 3: Researcher development awards

This form of funding can cover a range of circumstances, but in all cases the routine data collection would be from individual recipients of awards. Many funders have PhD fellowships schemes, but the MRC and other funders also have a range of post-doctoral fellowships. ~~The a~~Analysis for the North Thames region showed that whilst research training as indicated by PhDs would be the single most important aspect of payback to assess at this level, the full range of payback categories and indicators could also be appropriate. This is because some projects being undertaken by PhD Fellows can lead not only to publications but can have an impact on policy and practice.

A pilot evaluation of MRC research training schemes included issues such as the quality of training, and most importantly the pattern of employment and the extent to which recipients were contributing to, and leading, research and 'the extent to which MRC training and development awards had contributed positively to the career development of recipients' (MRC, 1996). In addition to indicators based on employment, and on continued involvement in research, some of the indicators used and recommended for future use involved assessing the publications record of former fellows. There were no indicators used or proposed that related to the impact of the research undertaken. It is also important to note that it was proposed that exercises of this kind should concentrate on individuals who had completed awards at least six years prior to the assessment.

Conclusions

This report is primarily concerned with showing how the benefits from a funding agency's research could be presented. We are proposing that payback be assessed by separately considering three main streams and various sub-levels within them, but this chapter has outlined a number of complications that arise when trying to do so. In the US the NIH is proposing that its GPRA report should be at full agency level. An advantage of operating at agency level is that it allows a portfolio approach to be used to explore the extent to which the programme as a whole shows an appropriate balance of risk, and of time-scales etc.

Nevertheless, data on various streams would be needed if the monitoring is to be used for the management purposes discussed earlier.

CHAPTER 6 : PROPOSED COMPREHENSIVE ROUTINE MONITORING SYSTEM

The purpose of this chapter is to describe a system that could be used to monitor R&D routinely. The system is designed specifically for use by the DH/NHS to monitor R&D they have funded, but is also relevant to other funders. It is designed to provide information useful both for justifying the DH's investment in R&D and for managing the department's R&D budget. The information necessary for the latter comprises information that will inform resource allocation decisions by refining estimations of expected payback from different uses of funds, as well as information providing *ex post* insight into the outcomes of R&D. Achieving the dual purposes of justification and of management affects the information that needs to be collected. A system designed solely to demonstrate the overall use of R&D funds would require less detail on individual projects, at less frequent intervals.

This chapter draws together insights from previous chapters about the prerequisites for a valid monitoring system. The central insight is that the system's components must be relevant to the funder's objectives. The objectives should inform decisions about which of the dimensions of value outlined in Chapter 2 must be measured by the system. The DH is concerned with payback across all dimensions, particularly health-related benefits. The system must therefore be comprehensive in the sense that it is capable of measuring all relevant dimensions. The DH and NHS fund both basic and applied R&D, so the system must also be able to accommodate the differences between these types of R&D, as outlined in Chapter 4. The DH funds R&D in three main streams, projects, centres and individuals, so the system must be sufficiently flexible to measure outcomes from each of these streams and at appropriate levels within them, as outlined in Chapter 5. Any system must also be operationally feasible, and meet the criteria outlined at the end of Chapter 3. It must be relevant to objectives, be decision-relevant, be capable of getting accurate information, minimise unintended consequences, and have acceptable net costs.

Taking these factors into account, we again recommend the use of a multi-dimensional approach. The arguments rehearsed in Chapters 2-5 suggest that no single indicator will give adequate information about all dimensions of payback from all types of R&D. The system we propose for projects, centres and individuals is based on three sources of information. The first source is information gained from direct monitoring of researchers and their organisations. Some important categories of payback would not be adequately measured by using only information from this source, so we recommend that this be supplemented by a second source, external databases, and a third source incorporating peer review, user surveys, analysis of policy documents and case studies of selected

projects and individuals. Without the third source, a routine monitoring system will not deliver sufficient information on health benefits to provide information useful to the DH.

The first and second sections of this chapter summarises the proposed system. The first describes the methods we recommend be used to collect information. The second section sets out the types of payback that we recommend be the focus of the system, and the outcomes and outputs we recommend be used as measures. The third section outlines the key issues that must be resolved before the proposed system can be operationalised, and uses the criteria set out in Chapter 3.

Sources of information for a routine monitoring system

The instruments available for measuring the payback from R&D were described in Chapter 2 . We propose using information derived from a number of these sources. We propose a system with three broad components: first, information that is collected directly from those funded in all three streams of R&D activity; secondly, information derived from external databases; and thirdly, supplementary information collected from a sample of projects, centres, or people.

1. Information collected directly from those funded, for all R&D activity. We propose that a set of information be collected for each project, each research centre and institution, and each researcher holding a personal award. At least for projects and researchers this information can be obtained using questionnaires (possibly administered electronically). A postal questionnaire approach was piloted in the North Thames NHS Region by HERG, Brunel University, and showed that this is feasible, although without clear sanctions there was an incomplete response.

Further work is necessary to determine how often the information should be collected and analysed. The length of the optimal reporting period depends on the costs involved in making and analysing reports, the benefits to the DH of having interim information, and the length of time before outputs and outcomes accrue. This period might be longer for centres than projects, if they receive infrastructure funding for extended periods.

If information is to be collected electronically, it may be feasible to administer this as a continuously updated electronic database. In other words, researchers, organisations or award holders may supply information, as and when it is available, directly onto the database. (An approach similar to this is used by ESRC's REGARD system.) Researchers would need only to add material to the database if they had something positive to say, in other words it would not have 'no' boxes.

This part of the proposed system must be consistent with other reporting requirements and infrastructure. Information collection should be co-ordinated and perhaps combined with the interim and final reports currently required of DH funded projects. As will be suggested in Chapter 7, additional work should be undertaken to investigate the pros and cons of standardising final reports across all DH projects.

Similarly, there are existing systems for reviewing organisations. There is an annual reporting period for recipients of NHS R&D programme Budget One funding; there are quinquennial reviews for a number of centres funded by the Department of Health; and research institutions are currently required to file quarterly reports on their R&D projects to the National Research Register. It is important that any new system be consistent with fixed parts of the existing system. Ensuring that a new system is compatible with existing practice is one way of reducing costs to researchers, thereby increasing the likelihood of compliance.

2. *External databases.* We recommend that external databases be used to supplement information collected directly. The citation indices and the Wellcome Research Outputs Database include bibliometric information about publications and some information about funders. This information should initially be used to validate responses obtained directly from individuals. It has the advantages of being in standardised format and reducing the administrative burden on researchers but, as discussed in Chapter 2, the disadvantage of being incomplete. We need empirically to establish how much information is lost by using a database.

3. *Supplementary information,* in the form of user surveys, review by experts and peers, case studies and analysis of sources used in policy documents. This last involves tracking the inputs to policy documents, such as guidelines. It is not currently routinely




undertaken. As will be discussed in the next section, this information should be collected on a rolling sample of projects, centres and individuals. It is, however, vitally important that it be collected if we are to gain information on health and economic benefits, and if we are to be able to verify the direct returns made by researchers and organisations.

Outputs and outcomes to be measured

The sources of information sketched above can be used to measure payback in the dimensions of primary interest to the DH. As shown in Table 3, these sources can be used to measure a range of outputs and outcomes, which collectively give information relevant to payback in terms of contribution to knowledge, research capacity, influence on policy, health benefits and economic benefits. The outputs and outcomes we recommend that information be collected upon comprise:

- the number of and nature of publications;
- the number of and nature of publications in the so-called 'grey literature', defined as material not published in formal outlets, in other words publications without an ISBN or ISSN;
- the number of and nature of patents;
- the contributions R&D makes to future R&D;
- development of research capacity, including researchers skills, higher degrees, and equipment;
- for personal funding, information about career development;
- whether the R&D has informed policy decisions or documents, such as guidelines;
- health benefits of the various types described in Chapter 2 ;
- economic benefits.

Table 3 Outputs and outcomes by measurement vehicle

-  We recommend collecting this information
-  Important to collect, but difficult
-  Not recommended, but preferred method for this category of payback
- Possible to use this method as a source for this information
- D** Possible, but difficult to use this method as a source for this information

	Self Reporting (questionnaires or annual reports)	External Datasets	User survey	Policy document source analysis	Peer/expert review	Case studies
Publications	- [1]	- [2]			-	-
Grey literature	- [3]					-
Patents	-	-		-		-
Contributions to further research	-	- [4]	- [5]		- [6]	-
Higher degrees	- [7]					-
Research capacity [8]	-				- [9]	-
Utilisation of existing research	D				-	-
NHS Staff development	-					-
Dissemination activities	-		-			-
Impact on policy documents/guidelines	-		-	-		-
Other political benefits	D		-			-
Health benefits	-		- [10]		- [11]	-
Economic benefits	D				D [12]	D

[1] number of publications and, for each publication, data on the journal, type of publication (whether an article, an editorial, and so on), and number of collaborating authors.

[2] relevant information contained in the Science and Social Science Citation Indices, Medline, and Research Outputs Database.

[3] number and type of output.

[4] citations analysis

[5] as part of a wider survey of users

[6] review by research peers

[7] number and nature of higher degrees

[8] includes, for example, the development of researchers' skills and additional equipment

[9] review by research peers

[10] survey of users in the health-services

[11] review by experts in the health services, by research peers, and by representatives of patients.

[12] review by economists and health service professionals.

We suggest that information on all dimensions except economic benefits be collected routinely, directly from researchers and organisations. We have omitted several categories of payback since, although involving useful outputs or outcomes, they are either ambiguous or difficult to measure. We do not recommend that information on the following be routinely collected, but Table 3 indicates the source recommended if they are deemed important:

- utilisation of existing R&D;
- development of NHS staff educational benefits from participation in R&D;
- dissemination activities;
- other political benefits.

As noted in Chapters 2 and 3, there are a number of dimensions of payback on which researchers and their organisations are unlikely to have full information. Moreover, the information supplied by those directly involved in R&D may not be reliable. The regular direct returns should therefore routinely be supplemented with the following information, undertaken on a sample of projects or organisations:

- user surveys should be undertaken routinely, using relevant users, on a sample of projects or institutions, to assess the contribution to policy, to future R&D capacity and possibly to health. Users provide a vitally important perspective on the value of outcomes. Their views can be obtained by surveys or by inviting representatives to join review panels;
- analysis of the sources used in particular types of policy documents such as guidelines (perhaps those endorsed by NICE), should be routinely undertaken to determine which R&D projects or institutions have influenced policy (Grant 1999);
- expert review should be undertaken routinely, on a sample of projects or institutions, to assess the following:
 - contributions to future R&D (where the experts comprise research peers);
 - development of research capacity (using research peers);
 - health benefits (using health professionals, managers, social scientists, and patient representatives);
- case studies should be used if economic benefits are to be monitored. Case studies can also be employed, to supplement the above by providing additional, in depth, information on any of the outputs or outcomes. The importance of health benefits to the DH suggests that case studies, the best method for identifying health benefits, should be an integral part of the routine system.

As will be discussed below, these supplementary methods are not only a vital source of primary information, but they also provide a method for verifying information collected directly from researchers and their organisations.

It is feasible to collect this supplementary information, using user surveys, expert review, source analysis and case studies, only on a subset of projects or organisations. The choice of a sample will be constrained by practical considerations discussed in Chapters 2 and 3. First, the size of the sample will be constrained by cost, meaning that the value of additional information generated by these techniques must be assessed against its cost. Secondly, the methods are not universally applicable. The users of R&D, for example, are often not able to attribute particular health gains and R&D outputs to particular projects, suggesting that user consultation will not be able to be applied to all areas. Chapter 7 suggests that additional work be undertaken to determine the appropriate method for choosing which projects, programmes or individuals are subject to what type of additional analysis.

The structure of the system, in terms of the dimensions measured and the measures used can be applied to centres as well as to individual projects. However, we have not resolved the issue of whether a centre needs to relate all of its outputs to individual projects, or whether some can be presented as an aggregate for the institution. In principle, the same categories can also be applied to awards to individuals. However, there may be an argument for restricting the focus for individual awards to the first six categories of Table 3.

Key issues

The system must be evaluated against the criteria set out in Chapter 3. It must maximise relevance to objectives, decision relevance and accuracy, and minimise unintended consequences and net costs.

Relevance to objectives

The system sketched above has been designed to be relevant to the DH's objectives of funding R&D which generates value in a number of different dimensions. It has also been designed to accommodate the differences outlined in Chapter 4, between basic science and applied R&D, between commissioned and responsive R&D, and to be appropriate for R&D related to social care.




Decision relevant

Decision relevance requires that the right types of information be collected at timely intervals, and presented in a form consistent with decision making processes.

Whether the system is decision relevant depends in part on solving the problem of financial attribution. The DH may wish to identify the separate and additional impact its funding makes to payback from R&D. In this case it is necessary to find a means of apportioning funding to outputs and outcomes. As shown in Table 4, in this case information will be needed about the different funding streams contributing to each publication, each grey output, each higher degree and each patent. This information should (or could) be supplemented by an estimate of the proportion of each output able to be attributed to each funding source. This is an ambitious requirement, raising conceptual and practical problems. It is likely to be difficult for researchers to attribute proportions of each output to different sources, and it may be resented by researchers, undermining the criteria of having a system that encourages compliance. It may, however, be a necessary task. It could be made more palatable if it were required by all funders, and could in the long-term perhaps become information that was routinely noted in publications. We propose the feasibility and acceptability needs to be tested.

Decision relevance also requires that information be able to be analysed and presented in a manner consistent with funders' decision making processes. It is difficult to achieve this simultaneously with the first criterion, which requires a degree of comprehensiveness and therefore entails the collection of information from a number of disparate R&D activities using a variety of different measures. The system that is sketched in this chapter will generate a huge quantity of data, on a number of different dimensions. This in turn leads to significant problems when analysing and then presenting information useful for managing the R&D budget. The first problem is that, within each dimension of value, the outputs generated by different types of R&D activity may not be comparable. For example, as was discussed in Chapter 4, publication patterns vary between disciplines and different publications might have different value. The second problem arises when trying to aggregate the outputs and outcomes from all of the dimensions. How, for example, will instances of policy influence be combined with or compared against instances of health gain? A process,

Table 4 Financial attribution of outputs

-  We recommend collecting this information
-  Important to collect, but difficult
-  Not recommended, but preferred method for this category of payback
- Possible to use this method as a source for this information
- D** Possible, but difficult to use this method as a source for this information

	Questionnaires	External Datasets	User survey	Policy document source analysis	Peer/expert review	Case studies
Publications	-	-				
<i>funding sources</i>	-	D				
<i>% each output attributable to funders</i>	D					
Grey literature	-					
<i>funding sources</i>	-					
<i>% each output attributable to funders</i>	D					
Higher degrees	-					
<i>funding sources</i>	-					
<i>% each output attributable to funders</i>						
Patents	-	-		-		
<i>funding sources</i>	-					
<i>% each output attributable to funders</i>	D					

probably subjective, will have to be designed to bring together all the disparate types of information collected by the system. Unless these problems can be resolved, the information generated by the system may not be useful for the purposes of managing the budget. Resolving this problem is relevant to ensuring that the final criteria of acceptable net costs, discussed below, is met. The costs of collecting detailed information can be justified only if it is possible to find a means of using that information to realise the benefits of decision-relevance.

A funder such as the DH may wish to make its decisions based on potential rather than on realised payback. On the one hand, it can be argued that if the DH delivers R&D projects, it is not necessarily its responsibility to ensure that results are implemented. On the other hand, it can also be argued that no distinction should be drawn between R&D and

implementation: that failure to realise benefits may be a consequence of the way the R&D was designed or the results disseminated. The DH needs to consider whether research has been implemented if it is to truly judge the value of its investment.

Accuracy

Any system must aim to generate information that is sufficiently accurate for its intended use. Information collected directly from organisations might be inaccurate because of errors or because of deliberate manipulation. Errors can be minimised by appropriate system design, by resourcing those making returns so that they have time to collect accurate data, and by monitoring returns.

We recommend that external databases be used as a means of monitoring information provided by researchers or their organisations. This can be achieved by, for example, comparing reported publications with those appearing in a database. In addition, those undertaking periodic reviews can be asked to comment on claims made in questionnaires; and *ad hoc* case studies can be carried out to amplify and verify particular reports or on particular issues. The use of an electronic data-base, available for scrutiny by other researchers, might also help maintain veracity. Furthermore, internal validations built into such systems 'help prevent projects from entering most out-of-range data, thereby improving data accuracy over traditional survey methods' (NSF, 1998).

Consistency between the reporting requirements of different funders may promote accuracy. If we are to move to a system of annual reporting with multiple funders, it is advantageous if all funders use the same format. This would limit the possibility of researchers manipulating the system by multiple reporting of single outputs, as well as minimising their administrative burden.

R&D outputs and outcomes accrue gradually over time. This means that outputs and outcomes relating to a single project, programme or person are likely to be reported over several years. It is important that the system is capable of eliminating duplicates, and that returns are compatible. Direct entry to a database that the researcher can review will help to avoid duplicates. But we believe that data will also need to be 'cleaned'.

In general, the problem of attribution has to be resolved if a system is to generate accurate information. As previously discussed, the attribution of impact on policy or practice to specific projects by researchers is difficult. The complexity of this is increased the longer the time lags involved, since this increases the likelihood that any outcome will be the result of a number of research projects involving researchers possibly funded from a number of

different agencies. It may be hard even for expert reviewers to attribute impact to particular projects or institutions. Moreover, where individual projects do make an impact, the time lags are sometimes so great that it may be difficult to ensure that they are picked up in a routine system.

Whether the process is acceptable to and is taken seriously by researchers depends on whether proper checks and incentives have been built in. There may be an adverse reaction and non-compliance from researchers and institutions, if they object to the amount of information requested. The objection might relate to the costs involved in providing the information or to the principle of reporting in an area where they have previously been relatively autonomous. We anticipate particular problems in obtaining information enabling funding to be meaningfully apportioned to each output. Whilst previous work has not found evidence of a systematic tendency to overstate impacts, if researchers and their organisations believe that funding relates to the information they give, they have an incentive to exaggerate the impact of their work.

Minimise unintended consequences

Performance measurement may give rise to unintended consequences if the measures include 'perverse incentives' and behaviour changes in response to those incentives. The previous section outlined some perverse incentives with respect to information disclosure. There may also be perverse incentives with respect to R&D activities, or dissemination of results. Measuring, and rewarding, only R&D that generated publications in high JIF journals, for example, may create an incentive to shift R&D activity from specialised clinical applications to research likely to lead to publications in those journals. This may, from the perspective of the DH, be an undesirable, unintended consequence.

Acceptable net costs

The direct costs of implementing indicators are borne by the monitoring authority and by the researchers and their organisations themselves. The costs to the monitoring authority may be considerable: Project Hindsight took more than forty person-years (Sherwin and Isenson, 1966). Data collection is necessarily costly, posing problems for public agencies in a policy environment emphasising the importance of reducing organisational costs.

The direct costs to researchers and their organisations may also be considerable. The growing administrative burden on UK researchers has been noted (Millar, 1998). Different indicators have different costs, and they are distributed differently between researchers and auditors. Bibliometric indicators based on public datasets do not necessarily involve any

costs to researchers (except perhaps in checking the accuracy of the dataset). Case studies, surveys and appraisals do involve researchers, and depending on coverage are likely to be more costly for the monitoring body. The total costs of the system and their distribution between different participants are therefore both important. Measures which require high administrative input from researchers are unlikely to succeed.

The costs include not only the direct costs of carrying out performance measurement. There may also be an indirect cost if researchers become disaffected by onerous or apparently 'pointless' reporting requirements. It may, of course, be impossible to completely avoid disaffection, since one of the aims of a system of performance measurement is to impose a degree of accountability. It must nonetheless be realised that this may affect research output and, therefore, payback. The gains likely to accrue from better monitoring of payback must always be compared with these losses, as well as the direct costs incurred by the monitoring body.

Overall administrative burden on researchers can be minimised by making the system for measuring value consistent with other reporting requirements, such as final reports.

This criteria requires that costs be acceptable relative to benefits. In practice, gross costs are also important, given the priority given to minimising organisational costs. It is also important that these costs be compared to the benefits generated by the system: the system itself be subject to payback analysis. Areas of likely payback include: focusing decisions, contribution to knowledge of paybacks, defend funding.

Conclusions

This chapter has sketched a system that can be developed for use as a way of routinely monitoring DH R&D. We have assessed the proposals against key criteria, recognising that a balance must be struck between comprehensiveness, accuracy, incentives and costs. The level of detail proposed is necessary if the DH wishes to use the system for managing R&D as well as for justifying the overall budget. If the latter were the sole objective, the system would need less detail and would also be less costly. The overall value of this system must be empirically tested, in relation to its intended uses.

If this system is to achieve DH's managerial objectives, it is essential that a method for financial attribution be determined. A prerequisite to this may be co-ordination with other major funders, notably the Medical Research Council and Wellcome Trust, at least to ensure that the concept of financial attribution is legitimate. The need for financial attribution and the feasibility of achieving it are both identified in the next chapter as areas requiring additional research.

The system we propose is incomplete, in that a number of elements need further refinement or specification. Whether the system in fact meets our criteria will depend on finding and implementing an appropriate design. Several of the components requiring further analysis have been identified and are discussed in the next chapter (Chapter 7).

CHAPTER 7 : RESEARCH AND MONITORING

This report is primarily concerned with making substantive proposals for a routine monitoring system, basing the proposals on existing research and previous studies. It will be clear, however, that on some issues we feel that there is not sufficiently clear evidence to give us full confidence in the feasibility of our proposals. We therefore identify in this chapter feasibility testing that needs to be undertaken before embarking on full implementation. In the case of some items of data proposed, we have argued that until such data has been collected for a period, we cannot be sure how much additional value it will contribute. Those issues, where we recommend monitoring of their specific value, constitute a second group of studies that will need to be undertaken. The database will, in addition to providing information for the main purpose identified, provide a rich data-source for research on payback and we indicate here some of the issues that it can illuminate. A range of further issues needs to be examined which may in turn lead to subsequent modification of the monitoring system. Finally, we emphasise that the system is intended primarily to meet specific objectives and needs itself to be evaluated in use.

Issues for feasibility testing

- The feasibility of, and researcher compliance with, requests to attribute proportions of research outputs to separate funding agencies/streams.
- Analysis of the feasibility of standardising final reports, both those within the DH research ambit and those of other funding agencies, and of developing such a standard format as far as possible.
- Testing the feasibility and user-friendliness of a system of on-line recording of data on projects.
- Testing the acceptability of these proposals to other major relevant UK research funders, and establishing whether their requirements would be met within this system.
- Testing the implications of alternative data requirements from centres.

Issues for specific evaluation once a system is implemented

- Comparison of self reporting of all publications with external database analysis of publications. What is the incremental value of the former?
- Analysis of where best to focus those aspects of monitoring, for example the use of user surveys, that it is proposed are applied to a sample of cases.
- Detailed analysis of the extent to which the database makes it possible to make meaningful comparisons between the impacts of different funding streams.

Research opportunities provided by the database

- Analysis of the correlation between publication outputs and the various other categories of payback to see whether publications can be used in some way as an acceptably good proxy for other outcomes. There is some evidence on the relationship between publications and quality (Cave *et al*, 1997), but no real evidence exists for the relationship between publications and 'payback'.
- Study of the correlates of research success, investigating the relationship between input, and research process, on the one hand and the quality of research on the other.
- Study of whether there appears to be any unnecessary redundancy of data or key items missing from the system

Other important related research issues

- Can we identify and test the added value performance indicators suitable to measure collaboration?
- The value of 'network analysis', which is not currently included in our proposals, should be reviewed as a possible indicator of the necessary conditions for impact, once the findings of a current ESRC study have been published.
- Development of a more robust and comprehensive classification system of research types for bibliometric analysis. This might replace the categorisation of journals into 'Research Levels'. In addition it might be possible to develop a system more relevant to the NHS for 'weighting' journals, to be used instead of existing Journal Impact Factors.
- Does research funding attract leading clinician researchers to the UK, and if so should it be included as an additional category of payback?

Value in use

- Study of whether the monitoring system has aided the DH in justifying expenditure on R&D by providing evidence of its value.
- Study of whether the monitoring system has aided the DH in improving management of R&D resources, through a better understanding of differential payback from different types of research or modes of funding.
- Analysis of what have been the incentive or disincentive effects of the monitoring system, on the various types of researcher involved.

Final comments

We have indicated above that a number of specific issues of feasibility need to be addressed prior to trying to introduce a monitoring system. We also stress that its value should not then be taken for granted but should be tested in the light of the use that is made of the system.

We have previously pointed to the fact that the need for greater accountability, and for formal monitoring is internationally a growing requirement for the proper management of public resources. However, the sort of system envisaged here goes, we believe, beyond what has been attempted anywhere else. What is now needed are informed views as to whether all, or part, of the system proposed would appear to constitute a good investment of resources, in terms of both central management effort and 'imposition ' of tasks on researchers. If so, there may still be a question as to whether this is the right time to embark on such a system.

The 'Leeds Castle Workshop' will provide a first opportunity to obtain a wide range of views about these important questions. The participation in this workshop is at least indicative that these are issues that others see as important and warranting serious discussion. We hope that this paper will provide a basis for such discussion, and subsequent development of ideas.

REFERENCES

Adams J *et al* (1997) Benchmarking of the International Standing of Research in England. *Report of a Consultancy Study on Bibliometric Analysis, Centre for Policy Studies in Education*, University of Leeds.

Audit Commission (1998) *Accident and Emergency Services Follow Up*. London: Audit Commission.

Black N (1997) A National Strategy for Research and Development: Lessons from England. *Annual Review of Public Health*, vol 18, 485-505.

Black N and Davies S (1999) Where do UK health services researchers publish their findings? *Journal of the Royal Society of Medicine*, vol 92, 129-131

Bozeman B and Rogers J (1998) Strategic management of Government-sponsored R&D portfolios: lessons from Office of Basic Energy Sciences Projects. In APEC, (ed.) *Proceedings of the APEC Symposium on the Evaluation of S&T Programmes among APEC member economies*. Asia-Pacific Economic Co-operation, Singapore, 9-41.

Bryan S (1997) Evaluation of genetic testing services. In Hanney S and Buxton M (eds) *Assessing Payback from Department of Health Research and Development: Second Report, vol 2: Ten Further Case Studies*. Uxbridge: HERG, Brunel University.

Buxton M and Hanney S (1994) *Assessing payback from Department of Health Research & Development: Volume 1: The Main Report. Research Report No19*. Uxbridge: HERG, Brunel University.

Buxton M and Hanney S (1996) How can payback from health services research be assessed? *Journal of Health Services Research and Policy*, vol 1, 35-43.

Buxton M and Hanney S (1997) *Assessing payback from Department of Health Research & Development: Second Report: vol 1: Main Report. Research Report No24*. Uxbridge: HERG, Brunel University.

Buxton M, Hanney S, Packwood T, Roberts S and Youll, P (1999) *Assessing the benefits from North Thames Research and Development, Research Report No 25*. Uxbridge: HERG, Brunel University.

Cave M, Kogan M and Smith R (eds) (1990) *Output and Performance Measurement in Government*. London: Jessica Kingsley.

Cave M and Hanney S (1996) Assessment of Research Impact on Non-Academic Audiences. *Unpublished report to ESRC*.

Cave M, Hanney S, Henkel M and Kogan M (1997) *The Use of Performance Indicators in Higher Education*. (3rd Edition) London: Jessica Kingsley.

Chief Secretary to the Treasury (1998) *Public Services for the Future: Modernisation, Reform, Accountability. Comprehensive Spending Review: Public Service Agreements 1999-2002*. CM4181. London: The Stationery Office.

Claxton K (1998) The value of information in the economic evaluation of new health care technologies: An application to a policy model of Alzheimer's disease. *Medical Decision Making*, vol 18.

Committee on Science, Engineering and Public Policy (1999) *Evaluating Federal Research Programs*. Washington DC: National Academy Press.

Comroe JH and Dripps RD (1974) Ben Franklin and Open Heart Surgery. *Circulation Research*, vol 35, 661-669.

Comroe JH and Dripps RD (1976) Scientific Basis for the Support of Biomedical Science. *Science*, vol 192, 105-111.

Cozzens SE (1997) The knowledge pool: measurement challenges in evaluating fundamental research programs. *Evaluation and Program Planning*, vol 20, 77-89.

Culyer A (1996) Taking advantage of the new environment for research and development. In Baker, M. and Kirk, S. (eds.) *Research and Development for the NHS: evidence, evaluation and effectiveness*. Oxford: Radcliffe Medical Press.

Dasgupta P and David P (1994) Towards a new economics of science. *Research Policy*, vol 23, 487-521.

Department of Health (1992) *Review of the Role of Department of Health Funded Research Units, The Williams Report*. London: Department of Health.

Department of Health (1994) *A Wider Strategy for Research and Development Relating to Personal Social Services*. London: HMSO

Drummond M, Cooke J and Walley T (1996) Economic evaluation in health care decision making: evidence from the UK. *Centre for Health Economics, Discussion Paper 148*. York: University of York.

Elderidge K and South N (1998) Evidence-based practice: Slow-acting remedy. *Health Service Journal*, 24-5.

ESRC (1997) *Assessing Research Impact on Non-Academic Audiences: Invitation to Tender*. Swindon: ESRC.

Fears R and Poste G (1998) *Radicalism, Rationalising or Rationing – What does the UK want from research in the science base and health service?* Fifth Annual SmithKline Beecham Science Policy Symposium. Harlow: SmithKline Beecham Pharmaceuticals.

Ferguson BA *et al* (1998) *Evaluation of Research Programmes in Northern and Yorkshire Region*. Report for NHS Executive Northern and Yorkshire Regional Office.

Gambardella A (1992) Competitive advantages from in-house scientific research: the US pharmaceutical industry in the 1980s. *Research Policy*, vol 21, 391-407.

Gibbons M and Johnston R (1974) The roles of science in technological innovation. *Research Policy*, vol 3, 220-242.

Gibbons M and Georghiou L (1987) *Evaluation of Research - A Selection of Current Practices*. Paris: OECD.

Gordon M and Meadows A (1981) *The Dissemination of Findings of DHSS Funded Research*. Leicester: Primary Communications Research Centre, University of Leicester.

Grant J (1999) Evaluating the outcomes of biomedical research on health care. *Research Evaluation*, vol 8, 33-38.

Hanney S, Packwood T and Buxton M (forthcoming) Evaluating the benefits from health R&D centres: a categorisation, a model and examples of application. *Evaluation: The International Journal of Theory, Research and Practice*.

Henkel M and Kogan M (1981) *The DHSS funded research units: The process of review*. Uxbridge: Department of Government, Brunel University.

Henkel M (1994) PSSRU research on care management: Kent community project. In Buxton *et al*, *Assessing payback from Department of Health Research & Development: Volume 2: Eight Case Studies, HERG Research Report No. 19*. Uxbridge: HERG, Brunel University.

Herbertz H and Muller-Hill B (1995) Quality and efficiency of basic research in molecular biology - a bibliometric analysis of 13 excellent research institutes. *Research Policy*, vol 24, 959-979.

Hills P and Dale A (1995) Research and technology evaluation in the UK. *Research Evaluation*, vol 5, 35-44.

HSMU (1999) *Good Practice in NHS/Academic Links*. Report by the Joint Medical Advisory Committee to the UK higher education funding bodies. Health Services Management Unit, University of Manchester [www.niss.ac.uk/education].

Jacob R and McGregor M (1997) Assessing the impact of health technology assessment. *International Journal of Technology Assessment in Health Care*, vol 13, 68-80.

Keen (Forthcoming) *Securing Better Value for Money*. Oxford: OUP.

Kogan M and Henkel M (1983) *Government and Research: The Rothschild Experiment in a Government Department*. London: Heinemann.

Lewis S and Jones J (1990) Output and performance measurement in government departments. In Cave M, Kogan M and Smith R (eds). *Output and Performance Measurement in Government*. London: Jessica Kingsley.

Lewis G and Dawson G (1998) The effect of funding on the outputs of biomedical research. *Scientometrics*, vol 41, 17-27.

Mansfield E (1991) Academic research and industrial innovation. *Research Policy*, vol 20, 1-12.

McConaghy D (1999) Measuring (accountable) success – analysis of *The Government's Measures of Success*. *Public Money and Management*, vol 19 (4), 62-5.

MRC (1996) MRC Research Training Awards: Pilot Career Path Study. Unpublished report to MRC.

NAO (1997) *The Defence and Evaluation and Research Agency: Review of Performance*. HC 411 Session 1997-98. London: The Stationery Office.

Narin F, Olivastro D, and Stevens K (1994) Bibliometrics/theory, practice and problems. *Evaluation Review*, vol 18, 65-76.

NIH (1998) *Government Performance and Results Act: Annual Performance and Results Act: Annual Performance Plan FY1999 and FY2000*.

NSF (1998) *National Science Foundation: FY 1999 GPRA Performance Plan*. [<http://www.nsf.gov/pubs/1998/nsf99gprapp/start.htm>].

Office of Science and Technology and Programme for Policy Research in Engineering Science and Technology (1993) *Returns to Research and Development Spending*. London: HMSO.

Our Competitive Future. Department of Trade and Industry. London, 1998.

Pollitt C (1993) *Managerialism and the Public Services* (2nd Edition). Oxford: Blackwell.

Richardson A, Jackson C and Sykes W (1990) *Taking Research Seriously*. London: Department of Health.

Salter A and Martin B (1999) The Economic Benefits of Publicly Funded Basis Research: A Critical review. *SPRU Electronic Working Paper Series, Paper No. 34*. Brighton: SPRU, University of Sussex.

Scott A (1997) Designing incentives for GPs. A review of the literature on their preferences for pecuniary and non-pecuniary job characteristics. *HERU Discussion Paper*, University of Aberdeen.

Sherwin CW and Isenson RS (1966) *First Interim Report on Project Hindsight (summary)*. Washington DC: US Office of the Director of Defence Research and Engineering.

Sizer J (1979) Assessing institutional performance; an overview. *International Journal of Institutional Management in Higher Education*, vol 3, 49-77.

Spaapen J and Sylvian C (1994) *Societal Quality of Research*. London: Science Policy Support Group.

Supporting Research and Development in the NHS. 1994. London: HMSO.

Thomas P (1985) *The Aims and Outcomes of Social Policy Research*. London: Croom Helm.

Treasury (1990) *The Government's Expenditure Plans, 1990-91 to 1992-93*. (Public Expenditure White Paper) Cm. 56. London: HMSO.

Treasury Select Committee (1999) *Public Service Agreements, Seventh Report of Session 1998-99*. HC378. London: The Stationery Office.

Weisbrod B (1971) Costs and benefits of medical research. *Journal of Political Economy*, vol 79, 527-544.

Williams D and Rank AD (1998) Measuring the economic benefits of research and development: the current state of the art. In APEC (ed) *Proceedings of the APEC Symposium on the Evaluation of S&T Programmes among APEC member economies*. Asia-Pacific Economic Co-operation, Singapore, 52-77.

Wilson J (1989) *Bureaucracy: What Government Agencies do and Why They do it*. New York: Basic Books.

Youll P (1997) Evaluation of the costs and quality of residential care for adults with learning difficulties. In Hanney S and Buxton M (eds) *Assessing Payback from Department of Health*

Research and Development: Second report, vol 2: Ten Further Case Studies: Uxbridge:
HERG, Brunel University.

BOOKS, DISCUSSION PAPERS AND RESEARCH REPORTS

BOOKS

1. Packwood T, Keen J and Buxton M (1991)
Hospitals in Transition: The Resource Management Experiment
Open University Press, (ISBN 0 335 09950 5) (£12.99)

DISCUSSION PAPERS

(ISSN 0951-6042 for series)

1. Martin Buxton (1986)
Heart Transplantation in the UK: The Decision-Making Context of an Economic Evaluation (£3.00)
2. Martin Buxton, Joy Ashby and Moira O'Hanlon (1986, revised 1987)
Alternative Methods of Valuing Health States: A Comparative Analysis Based on an Empirical Study Using the Time Trade-Off Approach in Relation to Health States One Year After Treatment for Breast Cancer (,3.00)
3. Brian Ferguson and Martin Buxton (1986)
Analysing Nottingham Health Profile Data: Identifying Common Response Patterns (£3.00)
4. Brian Ferguson (1986)
Using Health Profile Information to Quality-Adjust Survival Data (£3.00)
5. Richard Bell (1987)
Modelling the Costs of Heart Transplants by Patient Characteristics: A Perspective on DRGs (£3.00)
6. Bernie O'Brien (1987)
Donor Kidney Sharing and Game Theory: Exchange is No Robbery (£3.00)
7. Mark Sculpher (1989)
An Examination of the Incremental Cost-Effectiveness of a Screening Programme for Disease Detection: The Case of Treatable Diabetic Retinopathy (£3.00)
8. Bernie O'Brien and Julia Rushby (1989)
Measures of Outcome in Cardiovascular Cost-Benefit Studies: A Critical Review (£3.00)
9. Jacqueline Brown (1990)
An Annotated Bibliography of the Economic Literature Concerning the Evaluation of the Screening for Cervical Cancer (£5.00)

10. Mark Sculpher and Martin Buxton (1993)
The Private Costs Incurred When Patients Visit Screening Clinics: The Cases of Screening for Breast Cancer and for Diabetic Retinopathy (£3.00)
11. Mark Sculpher (1993)
A Snip at the Price? A Review of the Economics of Minimal Access Surgery (£10.00)
12. Mark Sculpher, Michael Drummond, Martin Buxton (1995)
Economic Evaluation in Health Care Research and Development: Undertake it Early and Often (£5.00)
13. Rachel Elliott (1995)
Secondary Economic Analysis in Intensive Care: The Case of Selective Decontamination of the Digestive Tract (£7.50)
14. Andrew Briggs, Diana Gibb, Mark Sculpher, Sandra MacDonagh, Tony Ades (1996)
Antenatal Testing for HIV Infection. Developing a Model for Economic Evaluation (£10.00)
15. Karen Gerard, Jackie Brown, Katharine Johnston, Alison Morton (1996)
The UK Breast Screening Programme: Local Service Organisation and Delivery (£5.00)
16. Katharine Johnston, Karen Gerard, Alison Morton, Jackie Brown (1996)
NHS Costs for the Breast Screening Frequency and Age Trials (£7.50)
17. Stirling Bryan, Jackie Brown (1997)
Extrapolation of Cost-Effectiveness Information to Local Settings (£5.00)
18. Jonathan Karnon, Jackie Brown, Dagmar Zeuner, Andrew Briggs (1997)
Issues in the Economic Evaluation of Antenatal Screening Programmes for Genetic Disorders (£5.00)
19. Katharine Johnston, Karen Gerard, Jackie Brown (1997)
Generalising NHS Costs From the Breast Screening Frequency and Age Trials (£5.00)

RESEARCH REPORTS

1. Joy Ashby, Martin Buxton and Hugh Gravelle (1987)
National Health Services Costs of the Intervention Services at Guildford, Huddersfield, and Nottingham in the Trial of Early Detection of Breast Cancer (£10.00) ISBN 0 902215 80 9
2. Chris Blades (1987)
Methodologies for the Estimation of the Running Costs of NHS Capital Developments: A Review of Current Practice (£5.00) ISBN 0 902215 81 7
3. Brian Ferguson, Martin Buxton and Michael Drummond (1988)
Measuring and Valuing Health States Relating to Visual Impairment: A Review of Literature, Concepts and Methods (£5.00) ISBN 0 902215 82 5
4. Joy Ashby, Martin Buxton and Hugh Gravelle (1988)
Does a Breast Screening Service Affect the Clinical Activity of General Practitioners? (£7.00) ISBN 0 902215 83 3

5. Martin Buxton, Tim Packwood and Justin Keen (1989)
Resource Management: Process and Progress. Monitoring the Six Acute Hospital Pilot Sites (£10.00) ISBN 1 872166 00 8
6. Joy Ashby, Martin Buxton and Hugh Gravelle (1989)
What Costs do Women Meet for Early Detection and Diagnosis of Breast Problems? (£10.00) ISBN 1 872166 10 5
7. Martin Buxton, Bernie O'Brien and Julia Rushby (1990)
Cost-Effectiveness of the Automatic Implantable Cardioverter Defibrillator: A Preliminary Assessment (£5.00) ISBN 1 872166 40 7
8. The Special Medical Development Project Team (1990)
Diabetic Retinopathy Screening Project: Special Medical Development at Exeter, Oxford and Sheffield (£15.00) ISBN 1 872166 75 X
9. Mark Sculpher and Martin Buxton (1991)
Report of Phase I of the Medical Laser Technology Assessment (£12.50) ISBN 1 872166 26 1
10. Martin Buxton, Tim Packwood and Justin Keen (1991)
Final Report of the Brunel University Evaluation of Resource Management - Summary (£7.50) ISBN 1 872166 31 8
11. Mark Sculpher (1992)
Screening for Diabetic Retinopathy: A Review of Available Options (£10.00) ISBN 1 872166 81 4
12. Alison Buist (1992)
The Oregon Experiment: Combining Expert Opinion and Community Values to Set Health Care Priorities (£10.00) ISBN 1 872166 91 1
13. Susan Kerrison, Tim Packwood and Martin Buxton (1993)
Audit in Practice. A Study of a Year of Audit Meetings in General Medicine at Four Hospital Sites (£3.00) ISBN 1 872166 37 7
14. Mark Sculpher (1993)
Phase II Medical Laser Technology Assessment: A Preliminary Economic Evaluation of the Diode Laser in Ophthalmology (£10.00) ISBN 1 872166 32 6
15. Mark Sculpher (1993)
Phase II Medical Laser Technology Assessment: An Economic Evaluation of the Nd:YAG Laser in Gastroenterology (£10.00) ISBN 1 872166 47 4
16. Susan Kerrison, Tim Packwood and Martin Buxton (1994)
Review of Supra-Hospital Audit in Medical Specialties (£10.00) ISBN 1 872166 82 2

17. Mark Sculpher, Jonathan Michaels, Mike McKenna, Julia Minor (1994)
Phase II Medical Laser Technology Assessment. The Role of the Laser in Percutaneous Arterial Revascularisation: A Cost-Utility Analysis (£10.00) ISBN 1 872166 92 X
18. Mark Sculpher (1994)
Final Report: Phase II Medical Laser Technology Assessment (£7.50)
ISBN 1 872166 97 0
19. Martin Buxton, Steve Hanney (1994)
Assessing Payback from Department of Health Research and Development: Preliminary Report. Vol 1: The Main Report (£10.00) ISBN 1 872166 03 2
19. Martin Buxton, Rachel Elliott, Steve Hanney, Mary Henkel, Justin Keen, Mark Sculpher, Penny Youll (1994) *Assessing Payback from Department of Health Research and Development: Preliminary Report. Vol 2: Eight Case Studies* (£15.00)
ISBN 1 872166 04 0
20. Moira O'Hanlon, Martin Buxton (eds) on behalf of the EuroQol Group (1995)
EuroQol Plenary Meeting, London, October 1994. Conference Proceedings (£10.00)
ISBN 1 872166 23 7
21. Noreen Caine, Linda Sharples, Martin Buxton, Andrew Briggs, Mike McKenna (1995)
The Development of a Cost Model to Reflect the Main Clinical Events Post Cardiac Transplantation (£10.00) ISBN 1 872166 28 8
22. David Wonderling, Martin Buxton, Christine McDermott (1996)
The costs and cost-effectiveness of cardiovascular screening and intervention. Economic evaluation of the British family heart study and comparisons with other studies (£7.50) ISBN 1 872166 59 8
23. Mark Sculpher, Diana Gibb, Julie Ratcliffe, Trinh Duong, A E Ades (1997)
Modelling the costs of paediatric HIV infection and AIDS: comparison of infected children born to screened and unscreened mothers (£10.00) ISBN 1 872166 54 7
24. Martin Buxton, Steve Hanney (1997)
Assessing Payback from Department of Health Research and Development: Second Report. Vol 1: The Main Report (£10.00) ISBN 1 872166 74 1
24. Steve Hanney, Martin Buxton (eds) (1997)
Assessing Payback from Department of Health Research and Development: Second Report. Vol 2: Ten Further Case Studies (£15.00) ISBN 1 872166 79 2
25. Martin Buxton, Steve Hanney, Tim Packwood, Simon Roberts, Penny Youll (1999)
Assessing the Benefits from North Thames Research & Development (£15.00)
ISBN 1 902316 010
26. A. E. Ades, Mark Sculpher, Diana Gibb, Julie Racliffe, Ramyani Gupta (1999)
A cost-effectiveness analysis of antenatal HIV testing in the UK (£12.00)
ISBN 1 902316 05 3

ORDER FORM: BOOKS, DISCUSSION PAPERS AND RESEARCH REPORTS

NAME:

ADDRESS:

.....

DATE:

	Price	Quantity
Book 1	£12.99	
Discussion Paper 1	£3.00	
Discussion Paper 2	£3.00	
Discussion Paper 3	£3.00	
Discussion Paper 4	£3.00	
Discussion Paper 5	£3.00	
Discussion Paper 6	£3.00	
Discussion Paper 7	£3.00	
Discussion Paper 8	£3.00	
Discussion Paper 9	£5.00	
Discussion Paper 10	£3.00	
Discussion Paper 11	£10.00	
Discussion Paper 12	£5.00	
Discussion Paper 13	£7.50	
Discussion Paper 14	£10.00	
Discussion Paper 15	£5.00	
Discussion Paper 16	£7.50	
Discussion Paper 17	£5.00	
Discussion Paper 18	£5.00	
Discussion Paper 19	£5.00	

Research Report 1	£10.00	
Research Report 2	£5.00	
Research Report 3	£5.00	
Research Report 4	£7.00	
Research Report 5	£10.00	
Research Report 6	£10.00	
Research Report 7	£5.00	
Research Report 8	£15.00	
Research Report 9	£12.50	
Research Report 10	£7.50	
Research Report 11	£10.00	
Research Report 12	£10.00	
Research Report 13	£3.00	
Research Report 14	£10.00	
Research Report 15	£10.00	
Research Report 16	£10.00	
Research Report 17	£10.00	
Research Report 18	£7.50	
Research Report 19 (Vol 1)	£10.00	
Research Report 19 (Vol 2)	£15.00	
Research Report 20	£10.00	
Research Report 21	£10.00	
Research Report 22	£7.50	
Research Report 23	£10.00	
Research Report 24 (Vol 1)	£10.00	
Research Report 24 (Vol 2)	£15.00	
Research Report 25	£15.00	
Research Report 26	£15.00	

I enclose a cheque for £_____

made payable to BRUNEL UNIVERSITY.

Send order to: THE SECRETARY
HERG
BRUNEL UNIVERSITY
UXBRIDGE
MIDDLESEX UB8 3PH

The price per copy is indicated and includes postage and packing

	Price	Quantity

