SURGICAL RESEARCH AND DEVELOPMENT IN THE NHS – PROMOTION, MANAGEMENT AND EVALUATION

Based on papers delivered at the OHE Conference, London, 25 November 1997

Edited by Katharine Johnston and Jon Sussex

Office of Health Economics
12 Whitehall London SW1A 2DY

ISBN 1 899040 36 6

Printed by BSC Print Ltd, London.

Office of Health Economics

The Office of Health Economics (OHE) was founded in 1962. Its terms of reference are to:

- commission and undertake research on the economics of health and health care;
- collect and analyse health and health care data from the UK and other countries;
- disseminate the results of this work and stimulate discussion of them and their policy implications.

The OHE is supported by an annual grant from the Association of the British Pharmaceutical Industry and by sales of its publications, and welcomes financial support from other bodies interested in its work.

CONTENTS

LIST OF CONTRIBUTORS

Professor Al Aynsley-Green
Nuffield Professor of Child Health and Director of Clinical Research and Development, Institute of Child Health and Great Ormond Street Hospital for Children NHS Trust, 30 Guilford Street, London, WC1N 1EH

Professor Peter Bell
Professor of Surgery; Chairman of the Research Board and a Member of the Council of the Royal College of Surgeons of England; Head of Department, Department of Surgery, University of Leicester, Robert Kilpatrick Building, Leicester Royal Infirmary, Infirmary Square, Leicester, LE2 7LX

Professor Martin Buxton
Professor of Health Economics, Director, Health Economics Research Group, Brunel University, Uxbridge, Middlesex, UB8 3PH

Professor Ara Darzi
Professor of Surgery, Department of Minimal Access and Colorectal Surgery, St Mary's Hospital NHS Trust, Praed Street, London, W2 1NY

Professor John Farndon
Professor of Surgery, University of Bristol, Department of Surgery, Bristol Royal Infirmary, Bristol, BS2 8HW

Professor Sir Miles Irving
Chairman, The Newcastle upon Tyne Hospitals NHS Trust, Freeman Hospital, Freeman Road, Newcastle upon Tyne, NE7 7DN; and formerly Professor of Surgery and Director of the NHS Health Technology Assessment Programme

Professor John Gabbay
Director, The Wessex Institute for Health Research and Development, University of Southampton, Biomedical Sciences Building, Bassett Crescent East, Southampton, SO16 7PX

Katharine Johnston
Research Fellow, Health Economics Research Centre, Institute of Health Sciences, University of Oxford, Oxford, OX3 7LF

Professor Richard Lilford
Professor of Health Services Research, University of Birmingham, and Director of Research and Development, NHS Executive West Midlands Regional Office, Bartholomew House, 142 Hagley Road, Birmingham, B16 9PA

Professor John Primrose
Professor of Surgery, University of Southampton, University Surgical Unit, Southampton General Hospital, Tremona Road, Southampton, SO16 6YD

Jon Sussex
Associate Director, Office of Health Economics, 12 Whitehall, London, SW1A 2DY

Professor John Wallwork
Professor of Surgery, Department of Cardiothoracic Surgery, Papworth Hospital, Papworth Everard, Cambridgeshire, CB3 8RE

Dr Paul Watson
Director of Health Policy and Public Health, North Essex Health Authority (and formerly Director of Acute Services, Cambridge and Huntingdon Health Authority), 8 Collingwood Road, Witham, Essex, CM8 2TT

Dr Simon Wood
Clinical Director, Safety and Efficacy Register of New Interventional Procedures (SERNIP), Academy of Medical Royal Colleges, 1 Wimpole Street, London, W1M 8AE

H Brendan Devlin
Brendan Devlin, retired Consultant Surgeon, Lecturer in Clinical Surgery at the University of Newcastle, Secretary of the National Confidential Enquiry into Perioperative Deaths (NCEPOD), and prominent figure in the drive to improve surgical performance, died in December 1998.

Chapter 1
Introduction: surgical research and development in context

KATHARINE JOHNSTON

Background

The UK National Health Service (NHS) makes a substantial investment in research and development (R&D) in order to maintain and improve the services it delivers. The NHS R&D programme is aimed at providing new knowledge that is generalisable and for wider public dissemination. This entails not only innovation but also evaluation and diffusion of health care technologies. Encouraging appropriate R&D in the NHS presents a challenge to health policy, if it is to:

- stimulate innovation in new services and ways of delivering existing services, in areas consistent with the needs of the NHS and with government priorities;
- obtain value for money in the NHS;
- ensure the diffusion of cost-effective technologies and restrain the diffusion of other, or unevaluated, technologies; and
- ensure that the net gains from innovation and diffusion of technologies are maximised.

These concerns are relevant for R&D relating to all types of health technologies, whether surgical, medical, diagnostic or preventive. In order to address these concerns, there is a need both for evaluation of all health technologies and for a greater understanding of the incentives for the stakeholders involved in the innovation, promotion and adoption of technologies.

The 1997 and 1998 NHS White Papers for England, Scotland and Wales (Department of Health, 1997; Scottish Office, 1997; Welsh Office, 1998) and the 1998 consultation paper for Northern Ireland (Department of Health and Social Services, 1998) all define general principles for the whole NHS, and share the same strategic aims of partnership and high quality services. In the particular context of R&D, the emphasis on quality and improving clinical effectiveness is

the most relevant policy theme. Following the White Papers, the National Institute for Clinical Excellence (NICE) has been established (on 1 April 1999) for England and Wales to appraise new and existing health technologies and provide clinical guidelines and methods of clinical audit. Clinical governance, an initiative to assure and improve clinical standards at a local level throughout the NHS, has also been introduced. In England and Wales the new Commission for Health Improvement supports and oversees the quality of clinical governance and of clinical services. In Scotland the issues of effective care and of the quality of care are being addressed by the new Health Technology Board for Scotland (HTBS), which will evaluate and provide evidence on the cost effectiveness of all innovations in health care. At the time of writing, arrangements for Northern Ireland have yet to be confirmed, but it is certain that the emphasis on improving the quality of health services will apply with equal force there too.

All NHS R&D operates within this changing NHS environment and a growing culture of evaluation. There are specific features of surgical R&D, however, that result in its promotion, management and assessment being distinct from medical R&D. This book attempts to identify and discuss these surgical issues. The book is based on the papers delivered and discussed at the Office of Health Economics conference held in London on 25 November 1997, updated to reflect subsequent developments. The text provides a commentary on the key issues in surgical R&D written by experts in the field and brings together recent thinking on how surgical R&D may best be promoted, financed, managed, evaluated and used by the NHS in the future.

In order to set the issues in surgical R&D in context, this introductory chapter describes some of the key features of the NHS R&D programme and reviews the past and present mechanisms for funding it. The description of the institutional arrangements applies specifically to England but the principles presented apply throughout the UK NHS. The chapter then moves on to the differences between surgical and medical R&D, which is a recurring theme of this book.

NHS R&D

The NHS R&D programme

In April 1991, the Department of Health established the first comprehensive R&D strategy for the NHS. The aim was to allow the NHS Executive to commission R&D directly on behalf of the NHS (Department of Health, 1991). The objective of the strategy was to create a health service in which clinical and managerial decisions are based on sound information. The aim was therefore to ensure that research questions are directed towards issues that matter to the NHS, that they are addressed with rigour, and that the findings are disseminated and implemented throughout the service.

The NHS R&D programme in England is organised centrally by the Department of Health and NHS Executive headquarters, on the advice of the Central Research and Development Committee. The programme is implemented nationally by the Department and Executive, and regionally by the eight Regional Offices of the NHS Executive. A programme of health technology assessment is managed and developed centrally by the NHS Executive's National Co-ordinating Centre for Health Technology Assessment (established in 1996), as advised by a Standing Group on Health Technology.

Until recently, nine national priority R&D programmes were identified to target research towards key areas, namely:

- cancer;
- mental health;
- cardiovascular disease and stroke;
- the primary/secondary care interface;
- physical and complex disabilities;
- asthma;
- mother and child health;
- dentistry; and
- methods to promote the implementation of research.

These programme areas have now been abandoned and current and future NHS R&D is being organised within three broad programmes. One of the three broad programmes is Health Technology Assessment, as before. The remaining R&D work is grouped into either the Service Delivery and Organisation programme, or the New and Emerging Applications of Technology programme. These three broad programmes are generally referred to by their acronyms, respectively: HTA, SDO and NEAT.

Concerns with the NHS R&D programme

Some aspects of the NHS R&D programme have been criticised. Firstly, there is concern about the research prioritisation process. Since resources for R&D are limited, topics have to be prioritised so as to try and maximise the potential benefits. The intention of the NHS R&D strategy is to let managers and users determine the topics researched. The Department of Health's Central R&D Committee (CRDC) makes the final decisions on priorities for R&D after consultations with experts and in the light of key government policy statements. The CRDC contains representatives from the research community, the NHS and users of NHS services. However, some argue that NHS R&D priority setting has not been based on the views of all relevant NHS practitioners, as the CRDC is dominated by clinicians and academics (Black, 1997). Furthermore, since the R&D programme has until recently been organised into ten programmes (the nine specific programmes listed above plus HTA), it was possible that if a topic did not clearly fit into one of them then it would not be funded (Millar, 1998). The recent reorganisation of the structure of NHS R&D into the three broad-based programmes of HTA, SDO and NEAT may help to overcome some of the earlier criticism, however.

The results of any prioritisation process clearly depend on whose views have been sought and how those views have been used. The appropriate criteria for prioritising R&D are not an area of debate that is unique to the NHS. The prioritisation of R&D is a research topic in its own right and is an issue in all countries. Researchers in The Netherlands have recently explored the use of societal criteria for setting R&D priorities (Oortwijn *et al.*, 1998).

The second major concern expressed about the NHS R&D programme is that there is a shortage of adequately trained research staff available to run the programme. The Richards Report (CVCP, 1997) highlighted this issue for doctors. But the same shortage of research staff also exists for nurses, the professions allied to medicine, and social scientists in the health care field.

Thirdly, a common criticism levelled at the R&D programme is its perceived lack of influence on clinical practice. This implies a need for more weight to be given to development: to the 'D' in 'R&D'. An R&D programme can legitimately research implementation methods but cannot take responsibility for implementation, argue Black and Mays (1996). Related to this is a concern about the value of R&D itself and the difficulty of measuring it (Buxton and Hanney, 1996).

Given the relatively recent inception of the NHS R&D programme, it is difficult to draw any firm conclusions yet about its overall success or otherwise. But the fact that the R&D programme has identified the importance of issues such as dissemination and implementation could be seen as an early and important finding.

Funding mechanisms

Overall funding for medical and health-related research in England amounts to well over £3 billion per annum. Industry (pharmaceuticals and medical devices) funds about £2 billion; research charities provide £340 million; the Medical Research Council (MRC) £278 million; and the Higher Education Funding Council for England (HEFCE) funds £190 million of R&D at medical schools, (Culyer, 1998). Little of this large allocation of resources is spent on surgery R&D however, although around 10% of current MRC funded trials are surgical. NHS R&D funding adds another £491 million to the overall total of R&D spend and most surgical R&D is financed from this element. The NHS funds are intended to cover not only the costs of the NHS R&D programme but also the NHS costs of hosting R&D supported by external funders.

The current system of NHS R&D funding is the result of the 1994 Culyer review: Supporting Research and Development in the NHS

(NHS R&D Task Force, 1994). Before giving further details on this, however, it is useful first to revisit the previous funding arrangements in order to provide some background to the recent changes and to identify some of the key issues in funding R&D.

Service Increment for Teaching and Research (SIFTR)

In 1976, the NHS introduced the Service Increment for Teaching (SIFT) in teaching hospitals to cover the additional service costs they incurred as a result of providing facilities for teaching medical and dental students. SIFT was an allowance per student to protect these hospitals from undue withdrawal of resources as a result of the Resource Allocation Working Party (RAWP) process. RAWP defined a formula for distributing NHS resources around the country more equally than had been the case previously. This had the effect of switching resources away from London, which had a concentration of teaching hospitals, and towards the rest of the country. The estimate of excess cost per student was based on a study comparing the costs of teaching and non-teaching hospitals which found that 75% of the extra cost per case in teaching hospitals was attributable to teaching (Culyer *et al.*, 1978).

In 1985, the government set up a review of RAWP and SIFT. As part of this, the excess cost attributable to non-commercially funded research was investigated. The analysis compared the costs from 60 acute teaching hospitals with the costs from 198 acute but non-teaching hospitals but did not identify separately the contributions of teaching and research (Department of Health and Social Security, 1988). The proportion of excess hospital costs protected from the RAWP formula was increased from 75% to 100% of median excess cost per student, thereby incorporating much of the costs of research activity as well as of teaching. In 1991, the government extended SIFT explicitly to cover all of the excess costs of teaching and research, and SIFT was renamed SIFTR.

Concerns with SIFTR

The precise additional costs caused by teaching and research in individual hospitals are difficult to quantify and SIFTR payments were believed to be significantly below the actual costs of research. There

were several areas of debate and confusion surrounding SIFTR (Chantler, 1992). Firstly, the selection of hospitals in the study on which SIFTR was based was argued to be arbitrary. Secondly, there was concern that SIFTR was being used to fund some patient care costs that would have been incurred regardless of research or teaching activities; for example, some of the costs of caring for the patients enrolled in trials (NHS R&D Task Force, 1994). Thirdly, there were concerns about the way that SIFTR was allocated and that the allocation for research should not be distributed according to the number of students being taught. Finally, there were concerns regarding the perceived inequity of access to SIFTR since only teaching hospitals were eligible, meaning that hospitals performing research but without teaching programmes were ineligible (NHS R&D Task Force, 1994).

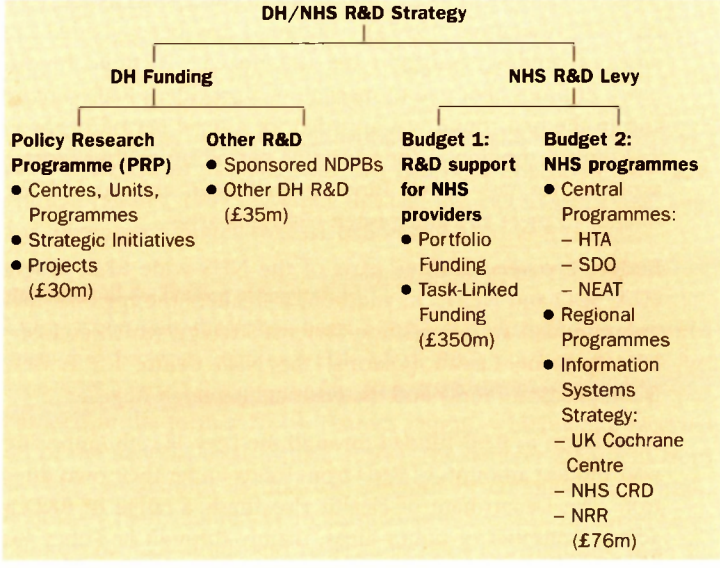
At the same time as doubts were being expressed about SIFTR, further concerns were being voiced about the impact of changes in NHS organisation on the funding and success of research. The introduction of the NHS internal market in 1991, following the NHS and Community Care Act (1990), separated the provision and purchase of health care. This separation was argued to have caused additional pressures on the funding of research (Drummond *et al.*, 1992). The concern was that health care providers would find it increasingly difficult to give adequate weight to the long term benefits of research in the face of short term pressures to demonstrate value for money to their purchasers (Parliamentary Office of Science and Technology, 1994). Research funders reported that the internal market was leading to problems in covering the additional costs of R&D and that some clinical trials were experiencing difficulties (Parliamentary Office of Science and Technology, 1994; Smyth *et al.*, 1994).

The 'Culyer Reforms'

These concerns led to the setting up of a NHS R&D Task Force in November of that year, chaired by Professor Tony Culyer. There was no question of the NHS abandoning the funding of R&D. Rather, the remit of the task force was to review the ways in which the NHS funded its own R&D and supported that funded by others. The aim was to make R&D funding more accountable and to make the process more explicit.

The Task Force reported in 1994 and one of its key recommendations was the setting up in April 1996 of a single national budget, the NHS R&D Levy, to fund R&D. By imposing a levy on all NHS purchasers of health care, a national budget is raised which provides a financial incentive for managers and clinicians to both undertake and participate in research. The Culyer Report's recommendations were intended not only to stimulate research but also to create a contestable research market in which providers would have to compete for research funding (NHS R&D Task Force, 1994).

Figure 1.1 **NHS R&D funding streams, 1998**



Notes:
CRD – Centre for Reviews and Dissemination
DH – Department of Health
HTA – Health Technology Assessment
NDPBs – Non-Departmental Public Bodies
NEAT – New and Emerging Applications of Technology
NRR – National Research Register
SDO – Service Delivery and Organisation

Source: Culyer, 1998, Figure 1.

An overview of the structure of NHS R&D funding in England that resulted from the Culyer reforms is presented in Figure 1.1. The R&D Levy is divided into two budgets: Budget 1 is used to support providers (NHS Trusts, GPs – primary care physicians contracted to the NHS – and charitable hospitals) undertaking R&D activity. Budget 2 is used to fund a NHS-wide programme to meet R&D priorities identified by the NHS Executive. The total sum of the NHS R&D Levy in 1997/98 was approximately £426 million, with Budget 1 being £350 million and Budget 2 £76 million (Culyer, 1998).

Budget 1 is designed to meet the costs that providers incur as result of being involved in externally funded R&D and is allocated on the basis of competitive bids, judged against published assessment criteria (NHS Executive, 1997a). Budget 1 has been divided into two types of funding: portfolio and task-linked. Portfolio funding is a block of funds allocated to meet all of a provider's R&D support costs and in the first round was awarded for a fixed period of three years. Task-linked funding is for a variable period up to four years. Those applying for task-linked funding have been able to bid for extra finance to meet any unexpected costs that arise.

Budget 2 covers all three parts of the NHS-wide R&D programme: HTA, SDO and NEAT. In addition, it finances the regional R&D programmes and an Information Systems Strategy, which consists of the UK Cochrane Centre (Oxford) the NHS Centre for Reviews and Dissemination (York) and the National Research Register.

In addition to R&D funded through the Levy, Health Authorities support modest amounts of R&D from funds under their own direct control. The Department of Health also funds a range of R&D activity across a number of policy areas, mainly through its Policy Research Programme, which aims to provide a knowledge base for Ministers and officials covering health and social policy. The Department of Health also funds some of the research undertaken by sponsored Non-Departmental Public Bodies such as the Health Education Authority and the UK Transplant Support Services Authority.

The implementation of R&D support

The initial implementation phase of the Culyer Report brought together existing central and regional budgets to create the NHS R&D Levy. At the end of May 1996, NHS hospital and community Trusts completed a declaration of their R&D activities and of the costs they incurred in supporting R&D. They were asked to identify the so-called 'service support costs' of supporting R&D funded by external organisations, such as the MRC. These consist of the extra patient care costs arising from R&D, such as the costs of extra tests or longer inpatient spells. Cross- subsidies between the R&D Levy and patient care were then, in principle, removed so that the actual costs of patient care were reflected in Trusts' contract prices charged to health care purchasers. The R&D Levy for the following year, 1997/98, was then based on the declared costs to the NHS of its R&D. For Budget 1, a competition and assessment bidding round took place and allocations for 1998/99 were announced at the end of 1997. Some bidders received modest increases in their support funding but some suffered a reduction in their allocations, partly as a result of the cut in the total R&D that had been imposed by the government (Swales, 1997). Despite this cut, the allocations of R&D support funding committed over £1 billion to R&D over three years.

Limitations of the funding arrangements

Critics argue that there has been a lack of redistribution between Trusts and between regions. Half of all R&D funding goes to just seven NHS Trusts. Geographically, NHS R&D funding is highly concentrated in the former North Thames region, which alone received half of the national total (Millar, 1998). The counter-argument from the R&D programme is that the strength of competition is high, resulting in many bidders being unsuccessful and that the current distribution of R&D funds still reflects previous distributions, with its perceived bias towards London and the South East. Although ultimately the aim is to remove any distribution bias, the allocation process has deliberately attempted to provide some stability for providers previously in receipt of large sums of R&D money (Swales, 1997).

Critics have also argued that research quality was not taken into account in the allocation process. Bids were supposed to include

quality indicators, but evaluating research quality is not an exact science and there are no universally accepted measures of it.

Although it is perhaps too early to judge whether the new R&D funding arrangements will be a success, they have already made the process more transparent by prompting efforts to explicitly identify the total amount spent on R&D. This process should become more reliable over time. In 1996 few NHS Trusts had in place information systems capable of identifying R&D costs as distinct from service and teaching costs.

The Clarke Review

In 1998 and 1999 the CRDC carried out a strategic review, under the chairmanship of Professor Michael Clarke, of priorities for NHS R&D and its funding (CRDC, 1999). This review confirmed the NHS R&D Levy as the preferred method of funding. However, from the financial year 2001/02 onwards, the allocations of money raised by the Levy will be re-labelled. All of the resources currently allocated as Levy Budget 2 (see Figure 1.1) plus some of the Levy Budget 1 (some of that for the NHS's own R&D, as opposed to support for partners' R&D) will become 'NHS Priorities on Needs R&D Funding'. The remainder of Budget 1 (i.e. that spent on supporting partners' R&D plus some of that used by the NHS for its R&D on its own account) will become 'NHS Support for Science'. (See Department of Health, 2000).

Surgical versus medical R&D

So far I have discussed the structure and funding of NHS R&D in general terms rather than specifically for surgery. This provides background to the system within which surgical R&D operates. The remainder of this chapter and of the book compares and contrasts the features that make the promotion, management and assessment of surgical R&D distinct from medical R&D.

Defining the innovation

The nature and definition of a health care innovation may differ between surgical and medical R&D. One of the ways in which surgi-

cal innovation differs from medical is the greater breadth of technological change that surgical R&D may encompass. For example, surgical innovation can be a new surgical procedure, a new device, a new pattern of service delivery or a combination of procedures, drugs and devices. There is a spectrum of scale of change in surgical innovation which at one end may include modifying an existing surgical technique and at the other end may include developing, for example, xeno-transplantation. Another potentially important difference between surgical and medical research is the greater role for the clinician in surgery: a medicine is a substance but surgery is a combination of a technique and the clinician that applies it to any given patient. The precise definition of the innovation has implications for the promotion and assessment of surgical R&D and this is discussed further by Professor Buxton in chapter 2.

Stakeholders in the innovation and diffusion process

As well as the surgeon and surgical team, there are other stakeholders with incentives in the innovation and diffusion of new surgical technologies. Unlike medical R&D, there is a lesser role for industry, unless the surgical innovation involves a device. The institutional arrangements of the hospital in question may introduce additional stakeholders with their own incentives, such as managers. The incentives held by the various stakeholders for surgical R&D are likely to be different than for medical R&D and the direction of impact of their incentives on the innovation, evaluation and diffusion of surgical R&D is discussed further by Professor Buxton (chapter 2).

Learning

For all types of health technology, there are issues surrounding the process by which clinicians, patients and researchers learn about innovations. In surgery, the initial learning often takes the form of learning by (supervised) doing. The surgeon practises a technique on patients, rather than learning through evaluative studies. Learning in surgery, as in other fields, is a continual process, since surgeons need to remain proficient with old skills while keeping up to date with newer ones.

After the initial learning process has taken place, evaluation is required but it is often difficult to determine the point in the learning curve at which a randomised controlled trial (RCT), or other means of evaluation, should take place. A balance has to be struck between performing a trial which may take several years before the results are established, but which has high scientific standards, and providing timely evidence for the NHS. This is an issue for all health technologies and presents a real problem for the design of evaluative studies. Professor Wallwork (chapter 3) gives examples of how this issue has been dealt with in practice.

Performing trials

Beyond the difficulty of determining the best timing for an evaluation, there are particular problems associated with performing trials in surgery. Practice variation from surgeon to surgeon leads to problems in trying to perform trials but there are additional problems of randomising patients. Examples of setting up surgical trials and of how some of these problems have been handled are discussed, with reference to case studies, by Professor Wallwork (chapter 3). He identifies several additional barriers to surgical research. For example, the evaluation of life saving but low volume-high unit cost surgical technologies, such as left ventricular assist devices, is difficult because of both a low recruitment rate to trials and problems in getting the high cost device funded. The problems associated with randomising in surgical trials imply that other types of evidence than RCTs may be appropriate for surgery. Brendan Devlin's argument for acceptance of other study designs is set out in the discussion to chapter 3.

Timing of research evaluation

Professor Wallwork argues that with respect to the devices industry there is a need for mechanisms to be introduced to encourage dialogue and early discussion between health departments, government agencies and device manufacturers, so that proper evaluation can be built into the development of the device. Common to surgical and medical R&D is the relationship of the timing of evaluation to its funding. If the research commissioning process could be speeded up, the research could be more relevant and timely.

Funding and priorities

A common complaint from the surgical research community is that surgery loses out to medical research, such as genetics, in terms of the total research funding it receives. Although there is no separate funding programme for surgery within the NHS R&D programme, 15% of the projects commissioned to date by the HTA programme (within Budget 2 – see Figure 1.1) relate directly to surgery. In chapter 4, Professor Gabbay discusses the process by which research priorities are set at a national level by the National Co-ordinating Centre for Health Technology Assessment. He also describes the work of the regional Development and Evaluation Committee (DECs) which, since April 1999, has been taken over by NICE. He describes a system of horizon scanning whereby new technologies likely to have a major impact on the NHS in the future are identified early in order to have time to perform the research and evaluation before any potential problems arise.

Dissemination

Although, as discussed above, there is an apparent lack of dissemination and implementation of the results of R&D, this problem has been recognised and there have been recent developments in both surgery and medical R&D to enhance dissemination. The NHS Executive produces circulars on clinical effectiveness across a range of health technologies. The Royal College of Surgeons of England and the Academy of Medical Royal Colleges have been involved in an initiative to promote dissemination of surgical research and have set up the Safety and Efficacy Register for New Interventional Procedures (SERNIP) to provide monitoring and rapid evaluation of new interventional procedures. Simon Wood describes SERNIP in more detail in his discussion of chapter 4.

The new National Institute for Clinical Excellence (NICE) will have a major part to play both in reviewing the cost-effectiveness of some surgical treatments and in producing and publicising clinical guidelines which encourage use of the most cost-effective procedures. Compliance with the new clinical governance framework being applied to health care providers in the NHS will require uptake of good surgical practice.

Delivering high quality R&D

SERNIP monitors the outcomes of surgical R&D rather than the processes that are required to deliver high quality R&D. In chapter 5, Professor Aynsley-Green argues that for all R&D in the NHS the organisational procedures which support R&D and the philosophies which underpin it have to be rigorously reviewed and assessed so that first rate R&D can be delivered. He discusses how these key, but often neglected, issues have been addressed in practice at the Institute of Child Health and Great Ormond Street Hospital. There, R&D structures have been reorganised and partnerships have been built between the university research institute and the NHS Trust.

The role of health care purchasers in R&D

With the recent changes to NHS R&D funding discussed above, there is a clearer recognition of who incurs the R&D costs and the service and treatment costs associated with R&D. The aim of the NHS R&D strategy is to create a health service in which clinical and managerial decisions are based on sound information but the issue of whether purchasers should implement the results of R&D and pay for patients to have access to new treatments is contentious. These issues are addressed from the perspective of a health authority by Dr. Watson in chapter 6. Technologies should only be introduced after appropriate consideration of the evidence, and the work of NICE should be of great assistance to purchasers in that regard. However, NICE will only be looking at a subset of the new treatments becoming available over time and even then appears likely to leave the final decision to local purchasers about whether, and to what extent, to make new services available locally.

Thus difficult decisions will remain for health authorities, Primary Care Groups and Primary Care Trusts; certainly in respect of the large majority of treatments not appraised by NICE and perhaps even in respect of those that NICE has looked at. The difficulty arises because new technology may increase costs either as a result of an increase in the unit cost of a treatment or as a result of higher volumes being treated because of a lowering of the threshold of treatment or the development of therapies for previously untreatable conditions. Dr. Watson explains the criteria used by health authorities when they

accept or refuse new technologies and argues that a managed process is required for the introduction of new technologies. The same arguments apply to the new Primary Care Groups and Primary Care Trusts with as much force as they do to the health authorities.

Research training

For both surgical and medical R&D, research training for the professions is crucial in order both to promote research as a valuable and important activity within the professions and to ensure high quality in its practice. A barrier in the past has been the lack of training specifically for research, and the competing demands for surgeons' time. In chapter 7, Professor Bell explains the specific mechanisms that have been introduced by the Royal College of Surgeons to formalise the training process and to widen the opportunities for surgeons to be trained in research skills.

Policy implications

In chapter 8, Professor Irving summarises the key issues raised throughout the book and draws out the policy implications. He argues that greater rewards and more time for research are needed. The opportunities available to surgeons to get involved in R&D are, he argues, growing, particularly with the introduction of the SDO programme which will have funds available to commission research into surgical services.

The central theme of the OHE conference and hence of this book is the features that make the promotion, management and evaluation of surgical R&D different from medical R&D. Specifically, the contributors have attempted to identify the factors that serve either to promote or impede high quality surgical R&D. The aim of this book is to stimulate debate and contribute to a better understanding of the critical issues in surgical R&D. As the NHS moves into a new phase of organisation, a re-examination of the key issues in surgical R&D is opportune.

REFERENCES

Black N (1997). A national research strategy for research and development: lessons from England. *Annual Review of Public Health* 18: 485-505.

Black N, Mays N (1996). What is 'development'? *Journal of Health Services Research and Policy* 1: 183-184.

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

Central Research and Development Committee of the Department of Health (CRDC) (1999). *Strategic Review of the NHS R&D Levy.* Final Report, October 1999. London: Department of Health.

Chantler C (1992). Service increment for teaching and research. Still problems concerning its definition and use. *BMJ* 305: 71-72.

Committee of Vice-Chancellor and Principals of the Universities of the United Kingdom (CVCP) (1997). *Clinical Academic Careers. Report of an Independent Task Force.* London.

Culyer AJ (1998). *Economics and Public Policy – NHS Research and Development as a Public Good.* Discussion Paper 163, Centre for Health Economics, University of York, December 1998.

Culyer AJ, Wiseman J, Drummond M, West P (1978). What accounts for the higher costs of teaching hospitals? *Social and Economic Administration* 12: 20-30.

Department of Health (1991). *Research for Health: a Research and Development Strategy for the NHS.* London, Department of Health.

Department of Health (1997). *The NHS: Modern, Dependable.* London, The Stationery Office.

Department of Health (2000). *Research and Development for a First Class Service – R&D Funding in the New NHS.* Leeds: Department of Health.

Department of Health and Social Security (1988). *Review of the Resource Allocation Working Party Formula: Final Report by the NHS Management Board.* London, DHSS.

Department of Health and Social Services (1998). *Fit for the Future.* Belfast, Department of Health and Social Services.

Drummond MF, Crump BJ, Little VA (1992). Funding research and development in the NHS. *Lancet* 339: 230-231.

Millar B (1998). Failing the acid test. *Health Service Journal* 26 March: 24-27.

NHS Executive (1997a). *R&D Support Funding for NHS Providers from 1998/9, Invitation to Bid.* NHS Executive, Catalogue number 97CC0057.

NHS Executive (1997b). HSG (97) 32.

NHS R&D Task Force (1994). *Supporting Research and Development in the NHS. A Report to the Minister of Health by a Research and Development Task Force Chaired by Professor Anthony Culyer.* London, HMSO.

Oortwijn WJ, Vondeling H, Bouter L (1998). The use of societal criteria in priority setting for health technology assessment in the Netherlands. Initial experiences and future challenges. *International Journal of Technology Assessment in Health Care* 14: 226-236.

Parliamentary Office of Science and Technology (1994). Research and the NHS reforms. *Science in Parliament* 51: 29-32.

Scottish Office Department of Health (1997). *Designed to Care: Renewing the National Health Service in Scotland.* Edinburgh, The Stationery Office.

Smyth JF, Mossman J, Hall R, Hepburn S, Pinkerton R, Richards M, Thatcher N, Box J (1994). Conducting clinical research in the new NHS: the model of cancer. *BMJ* 309: 457-461.

Swales J (1997). *R&D Support Funding – Allocations from April 1998* (Letter to NHS providers bidding for R&D support funding). Department of Health web site: www.open.gov.uk/doh/rdd1.htm.

Welsh Office (1998). *Putting Patients First.* Cardiff, The Stationery Office.

Acknowledgement

I would like to thank Bronwyn Croxson and Jon Sussex for helpful comments made on an earlier draft.

Chapter 2
Economic aspects of surgical research and development in the NHS: an overview

PROFESSOR MARTIN BUXTON

Introduction

'Health technology' has been defined as 'interventions used by those in the NHS to promote health, prevent and treat disease, and improve rehabilitation and long-term care' (NHS Executive, 1999). 'Health technology assessment' is the consideration of 'the effectiveness, appropriateness and cost of technologies' (NHS Executive, 1999). The important point to stress is that the definition of health technology assessment includes the clinical assessment of health technologies and is therefore about the effectiveness of technologies and not merely their broader social or cost impacts. In introducing the key issues facing the assessment of surgical technologies, I will focus on the particular characteristics of surgery that may make it different from other health technologies.

Is surgery different?

From the perspective of the promotion, management and assessment of surgical R&D, it is important to consider whether surgery is different from other categories of health technology. Does it face a different set of problems? If there are differences, which are the most important and what implications do they have for facilitating good surgical R&D?

On the face of it, surgical research is somewhat different from research comparing Pill A with Pill B. There are four key areas I will highlight. Firstly, there is a set of issues which relate to how to define and differentiate new surgical technologies. For example: when is something new, when is something different, when is it merely a modest variant of what other people have been doing? Secondly, there is a big problem about how surgeons learn to do surgery and what that then implies, not only in terms of when to evaluate but also how to fund surgical research. Thirdly, there are some well recognised difficulties in undertaking RCTs in surgery, perhaps not insurmountable difficulties, but difficulties which remain important in

explaining the paucity of surgical trials. Finally, the incentives to innovate, to evaluate and to diffuse new surgical technologies may differ from those for other categories of health care technology. These four key areas are now discussed.

Defining and differentiating new surgical technologies

One of the key questions is what are the characteristics of a 'new' surgical technology? When do variations in the way an individual or institution undertakes a procedure make it a 'new' technology? Even with traditional surgical techniques, there is quite considerable variation between individual surgeons in precisely how a procedure is performed. There is therefore a fundamental question as to how 'new technology' is defined and when it is clear that a variation has become important enough to be considered and evaluated separately.

Linked to this is the idea of the extent to which one can talk about classes of interventions, as with, for example, drugs. In surgery, is it sensible to talk about minimally invasive hernia techniques as a class that can be compared with traditional ones? Or is there so much variation within them that it is meaningless to group them together? Arguably this a similar to the problem with medicines of whether, for example, we consider all ACE inhibitors to be of the same class or whether we need always to distinguish between individual ACE inhibitors.

Finally, there is a danger that when defining surgical technologies the focus is too much on the surgical technique itself. Yet the surgical intervention in its entirety is more than just the surgery and may include, for example, devices and drugs. This has important implications for what is evaluated and how the evaluation is undertaken. From the patient's point of view, what is of concern is the total treatment package from assessment through to discharge, and through to any subsequent complications arising, rather than precisely how the surgery is done for the hour or so that they are in theatre.

Learning in surgery

There are two types of learning in surgery which require distinction. The first is where the learning concerns developing a new technology. In this case, a pioneering surgeon learns as a developmental process,

trying to work out how to perform a new technique or how to change an existing one. The second is successor learning, where other surgeons learn how to perform what is to them a new technique but one which has already been precisely described or demonstrated.

The nature of learning relates back to the definition of new technology. If the surgical technology is the whole therapeutic process, then it is not just the surgeon who learns but the whole surgical team working with them. In the case of heart transplantation, an area of research that I have been involved in, some of the issues were not just about how the surgeon learned to do it, but also about how a whole team learned to deal with heart transplant patients, for example, how to nurse them.

There is a need to consider very carefully how the developmental learning process is funded. Going back to the example of the heart transplant programme, one of the criticisms that was levied frequently in discussing the issue about the timing of its evaluation, was that if no one is learning and developing the process, then it will never get to the point where the technique is cost-effective. Someone has to go through a process of learning, and improving the technique and they will need funding for that. On the other hand, it would be inappropriate to continue to provide R&D funding to the point where people are no longer really 'learning by doing' but rather are simply 'doing' while claiming that they are still learning or developing the technique.

R&D or *service?*

If we are in the development stage of a surgical technique, this then has to be judged not in terms of its benefits to the current patients (although one hopes that they do benefit) but in terms of its future 'payback', a terminology I have used elsewhere (Buxton and Hanney, 1996). This means the ability to generate future benefits by identifying good practice, or conversely by identifying inappropriate practice, that will create knowledge and so improve the way that patients are treated in the future. Patient services, on the other hand, have to be judged in terms of their cost-effectiveness in delivering patient benefits from the technology. Just as R&D and patient services need to be evaluated according to different criteria, so should the funding of these types of activities be kept separate.

Difficulties in undertaking randomised controlled trials in surgery

There are important problems in undertaking RCTs in surgery as against other types of health technologies such as medicines (Table 2.1).

Table 2.1 **Problems in undertaking RCTs in surgery**

- Surgeons
- Skill, training and experience
- Patient preferences: clear short-term differences
- Surgeon preference: lack of equipoise and bias
- Lack of blinding
- Placebo effects
- Multicentre variability and competitiveness

I have put the surgeons themselves at the top of the list of problems. Some of the characteristics that make a good researcher are not the same as those that make a good surgeon. As a patient, I am not sure that I want too much equipoise as the surgeon stands there ready to intervene. I would rather like to think that the surgeon knows what he is doing and that he is simply going to go into the theatre and do it. I would also want to make sure that I agreed with what he is proposing. I think there is an issue, however, that surgeons may not be temperamentally so well suited to the mindset required for R&D.

There are also the problems, that have been well documented by others, about how one takes into account in a RCT surgeons' differing degrees of skill, training and experience, and where they are on their learning curves for the specific procedure to be assessed. I do not think we have yet worked out quite how to handle these difficulties.[1]

1 Since the presentation was originally given, increasing attention has been directed at some of these methodological issues. For example, the concept of 'tracker trials' has been proposed by Lilford et al. (2000), and a review of methods for assessing learning curves, commissioned by the NHS HTA Methodology Programme, is soon to be published (for further details see the NHS R&D HTA Programme website www.hta.nhsweb.nhs.uk).

A further problem in performing surgical trials relates to patients' preferences and the implications of these for recruitment to trials. It can be quite problematic to get informed consent from patients to participate in trials of minimally invasive techniques compared with more traditional techniques, for example. One might explain to a patient: 'we don't know in terms of long term outcome whether treatment A is better than B, but we would like you take part in this trial'. But if the patient knows that treatment A is minimally invasive and he is going to be about and walking normally in a week, whereas treatment B is traditional and he will then need three weeks to recuperate and three months off work, then, quite reasonably, the patient will base his choice on the one difference that is known. It may be very difficult to find patients willing to consent to having the traditional surgery.

Surgeon preference can be a problem in some trials. There are examples where surgeons perform a long case series, publish the findings and say how good the technique is, and then are persuaded to do a randomised trial where they themselves are randomising between the technique that they have just argued is the best, and the one they have argued is inferior. As there is not going to be full blinding on either side, this really is a problem.

Then there is the issue of placebo effects. It is interesting to go back to some of the trials undertaken in the 1950s in order to illustrate this. One in particular was a study of 17 patients with angina who were randomly assigned to receive either mammary artery ligation, or a sham operation involving the same incision and exposure of the mammary artery, but in which the artery was not ligated (Cobb *et al.*, 1959). The study showed equal improvement in both groups. Of course, it was mainly used to argue that ligation did not work but an alternative interpretation could be that patients may feel better if something rather drastic is done to them, whatever that actually is.

Variability is practice between centres in a multi-centre trial is inevitable and difficult to standardise for. Competitiveness between centres may result in single centres undertaking trials with small numbers of patients and lack of statistical power. Multi-centre trials

are preferable but require collaboration and agreement about basic aspects of design including definition of the intervention and its comparator.

None of the factors relating to the ability to perform RCTs in surgery are insurmountable but nor can they necessarily be easily resolved. They should not be forgotten when trying to think about the manner in which surgical research has progressed, or why slower progress appears to have been made in surgical areas than in some others.

Incentives to innovate, evaluate and diffuse surgical techniques

Surgery differs from other heath technologies in terms of the principal incentives for the stakeholders to innovate, to evaluate and to adopt new technologies. The most important incentives are not necessarily economic or financial ones. Professional concern for patients figures large. Professional and institutional prestige are also very important. There may be peer pressure to adopt new technologies or to drop outdated ones, although it is difficult to assess how strong this pressure actually is. There are also financial and commercial incentives, and there may be (and perhaps could be more) managerial and political pressures, if these were deemed appropriate.

Different stakeholder groups are faced with different sets of incentives. The surgeon, or surgical team, has incentives to perform both NHS and private work. Approximately 13% of elective surgery is now in the private sector, or privately financed. It may well be that different incentives are operating for NHS work and private work, not least because the financial arrangements for each type of work are very different. The hospital or institution has different incentives, as does the third party payer, depending on whether the payer is a NHS commissioning body or a private insurer. Professional bodies also have their own incentives. A final interested group is private industry. Whilst surgery does not typically have patents associated with it, and unlike medicine is not owned by commercial interests, many of the devices and instrumentation associated with surgery do and are. For these there may be quite considerable commercial interests to try to encourage innovation and diffusion amongst surgeons of new devices and technology.

Figure 2.1 **An impressionistic view of incentives for R&D in surgery**

	To innovate	To evaluate	To adopt
Surgeon:			
NHS	✔	?	?
Private	✔	✘	✔
Hospital:			
NHS	?	?	✔
Private	✘	✘	✔
Third party payer:			
NHS	✘	✔	?
Private	✘	✔	?
Professional bodies	?	?	?
Manufacturers	✔	?	✔

Figure 2.1 summarises my, purely impressionistic, view of which stakeholders currently face incentives to innovate, to evaluate and to diffuse. There is not a lot of good evidence. In Figure 2.1, ticks imply a positive incentive; question marks suggest uncertainty; and crosses show a negative incentive. Whilst this may be over simplistic or even wrong, in the end it is these incentives that have to be understood in order to influence the innovation, evaluation and diffusion of surgical technologies.

Policy developments

Future policy developments in relation to surgical R&D may be informed by revisiting previous recommendations. A previous set of thinking about the costs and benefits of surgery led to the four main recommendations set out in Table 2.2. These recommendations come from a book by John Bunker *et al.* published in 1977. It seems rather disappointing in some ways that when you go back to that book, although some of the examples used are perhaps not the most up-to-date, many of the issues raised in it are still very much the ones we face today.

Table 2.2 **Four previous recommendations:**

- The need for further studies of surgical treatments
- Improving and extending our technical ability to do such studies
- The teaching of economic, social and epidemiological principles of medical care
- Improving the public's understanding of medical outcomes and costs

Despite this, there have been some important developments in the UK in the last 20 years. I have not gone back to the MRC's records and worked out how many trials there have been at different points in time but there does now seem to be more active support and encouragement within the MRC to support surgical trials. In the last few years, both the NHS R&D programme and, more particularly, the Standing Group on Health Technology have been established. Looking at all the projects that have been commissioned at this point on behalf of the Standing Group on Health Technology, about 15% appear to relate directly to surgery. (That is my estimation, and clearly it depends on where you draw the dividing lines.) A further development has been the establishment of SERNIP, discussed in more detail by Simon Wood in chapter 4 below. This is an important step forward, but perhaps one that needs still to be built on if it is really going to deliver some of the desired answers.

But despite all these things, despite some encouraging signs, despite the fact that the importance of these issues has been known for 20 years, a feeling remains that it is not actually getting any easier to conduct good surgical trials.

REFERENCES

Bunker JP, Barnes BA, Mosteller F (eds.) (1977). *Costs, Risks and Benefits of Surgery*. Oxford University Press, New York

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

Cobb LA, Thomas GI, Dillard DH, Merendino KA, Bruce RA (1959). An evaluation of internal-mammary-artery ligation by a double-blind technic. *The New England Journal of Medicine*; Vol. 260, No.22, May 28, 1959, pp. 1115-1118.

Lilford RJ, Braunholtz DA, Greenhalgh R, Edwards SJL (2000). Trials and fast changing technologies: the case for tracker studies. *BMJ*, 320: 43-46.

NHS Executive (1999). *The Annual Report of the NHS Health Technology Assessment Programme* 1999. London, Department of Health.

Acknowledgement

I would like to acknowledge the contribution of Katharine Johnston, upon whose ideas I have drawn.

Discussion

PROFESSOR JOHN FARNDON

When I was invited to discuss Professor Buxton's paper I could not help but ask the question, 'Why is surgery the focus of this attention?' Why the interest in surgical disciplines, when I believe that surgeons have been the leaders in audit of our activity? My predecessor in the University of Bristol, Professor H Groves, first wrote a paper on the need to audit surgical activity, to measure what we were doing and what its outcome was, 60 years ago.

I believe we do as well as any other group in trials and research, but I think that there may be some fundamental limitations still applying to surgical research. Amongst the medical profession, however, there have been examples of biased, uninformed opinion, such as that propagated in The Lancet (Horton, 1996). At the same time there is sometimes triumphalism, for example, the BMJ cynically praised the surgical profession for producing a controlled trial which looked at octreotide in the treatment of portal hypertension (Jenkins *et al.*, 1997). Surgery attracts attention, more than some other areas with smaller evidence bases, because its end results are often well defined and relatively easy to measure.

Martin Buxton has highlighted the difficulties encountered when undertaking RCTs in surgery, and I agree with him that surgeons are one of these problems. The concept of the learning curve is important. It can be very rapid with a surgical innovation. One of the frustrating things that we have had to address on the Acute Sector Panel of the Health Technology Assessment Programme, is spotting the emergence of a new technology and trying to capture and control it so that it is appropriately studied before being put into widespread use. Additional constraints on performing good surgical research are commercial pressures from industry and financial pressures resulting from the competitive elements between NHS Trusts; pressures that were perhaps not so strong a few years ago. Our technical ability to perform trials also depends on funding. Determining how more funding might be forthcoming through the Health Technology Assessment Programme and the MRC is therefore critical for the future of surgical research.

Turning to recommendations for the future, I am not sure that we need a greater flow of studies. As Editor of the British Journal of Surgery, I already receive about 1,400 manuscripts every year. So I am not sure that quantity is necessarily a problem. Maybe the sort of studies conducted is a problem, and this needs to be addressed. Then there is the question of which particular treatments to give research priority to.

A very important aspect is improving the public's understanding of medical outcomes and costs. An awful lot more needs to be done to educate the public at large about their own bodies and how they work. Related to this, we also have to learn how to influence and use the media when we are trying to interact with the public about our research effort.

SERNIP (Safety and Efficacy Register of New Interventional Procedures), discussed by Simon Wood later, in chapter 4, is an exciting development and we have got to support it and strengthen its activity. This is a group that wants to receive voluntary notification of new procedures, mainly in the surgical disciplines. It has a broad background committee, linkage to the Health Technology Assessment Programme, the MRC and the Medical Devices Agency, and is intended to inform health care purchasers and providers, and the Standing Group on Health Technology. I am sure that this development is an important way forward in addressing some of the difficulties faced by surgical research.

REFERENCES

Horton R (1996). Surgical research or comic opera: questions but few answers. *The Lancet*; 347: 984-985.

Jenkins SA, Baxter JN, Critchley M, Kingsnorth AN, Makin CA, Ellenbogen S, Grime JS, Love JG, Sutton R (1997). Randomised trial of octreotide for long term management of cirrhosis after variceal haemorrhage. *BMJ*, November 22; 315: 1338-1341.

Chapter 3
How do surgical innovation and diffusion occur in practice?

PROFESSOR JOHN WALLWORK

Introduction

The NHS R&D strategy and other pressures for changes have led to a concern that new surgical technologies be properly assessed before being introduced. In considering how best to develop policy to address this concern, it is useful to review the current issues in surgical research. The aim of this paper is to define some of these issues, illustrate them with case studies and propose a model for change in surgical research.

Surgical innovation

A new surgical technique goes through several development phases before it is diffused into wider practice (Table 3.1).

Once surgical innovation has taken place, surgeons have to learn and refine the techniques of perioperative care. The results are then audited and the techniques then further refined in an iterative process. Finally a new procedure is adopted and is generally within the capabilities of the average well trained surgeon. This process of introducing new surgical techniques is, however, not research. It is audit and there is clearly quite a difference between audit and research. The question is, at what stage in the development procedure is it appropriate to perform a trial to establish the efficacy and benefits of the procedure? I would suggest that the best time to begin proper eval-

Table 3.1 **Phases of development of new surgical techniques**

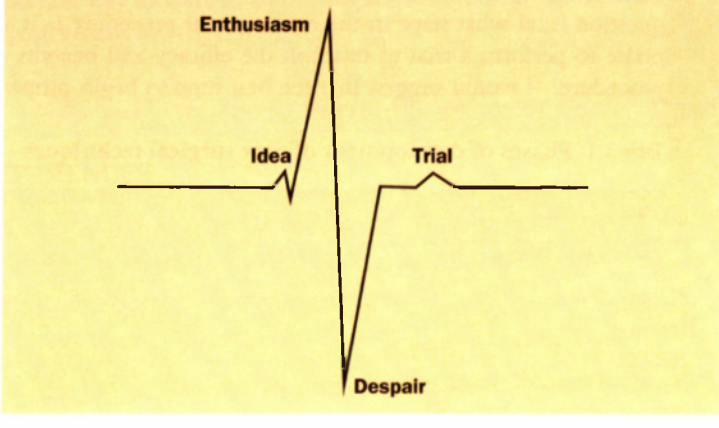
- Innovation
- Audit of results
- Refinement of technique and post operative care
- Adoption of technique within the capabilities of the 'average' well trained surgeon
- Appropriate trials and evaluation?

uation is after the initial audit process, that is, before it has diffused into general use. It should not be presumed, just because we can do something safely and well, that we *should* do it.

Currently there is often an enthusiastic innovator who develops a new technology. Then there is a band-wagon of others wanting to try it out. A surgeon attends a specialty meeting and listens to the experts who persuade him of the advantages of the new technique and the surgeon then goes off and uses it because he fears being left behind. Then at the next meeting there is a great pro and con debate: two luminaries of surgery stand up and debate whether a technique should be used. The conclusion of the debate is that a controlled trial should be performed but, more often than not, it never gets done.

Figure 3.1 illustrates the phases of learning and the diffusion of ideas. First, a new idea, followed by initial enthusiasm represented by the upstroke of the ECG where the thinking is 'it needs to be done to everybody.' After a while, despair sets in where the thinking has become 'one wouldn't do it to one's dog'. This is represented by the downside of the ECG. Then afterwards, in a more quiescent phase, some sensible idea of a trial occurs. We should probably not dampen

Figure 3.1 **'ECG' of a new surgical technique**



the initial enthusiasm but we could avoid the wild swings of enthusiasm and despair by appropriately timed surgical trials.

An example of a current development within cardiac surgery, using minimally invasive surgery, illustrates some of these points. Minimally invasive surgery has some potential advantages in reducing patient discomfort, shortening hospital stay, reducing rehabilitation, and reducing health care costs but these advantages have been promoted as facts before being proven. As a consequence of simply believing in these potential advantages, practice is changed, for example by sending patients home a day early, thereby biasing results.

The surgery involves a small incision being made. After the operation only a small surgical scar is visible and the patient is expected to go home after two days. The innovative feature is that the retractor and other equipment used are disposable. The growth in the development of disposable equipment is one of the pressures that industry places on surgical innovation and surgical trials. Industry wants to get its equipment onto the market as soon as possible and often does this by making the equipment disposable, without waiting to see whether it is of benefit or not.

Pressures and problems in performing surgical research

Horton's commentary in the Lancet (Horton, 1996) was mentioned earlier. Horton was essentially saying that innovation and surgical trials are done badly, that there are very few that have been done well, and he had some vague evidence to support this contention. He may not have been entirely right, but there was an element of truth in his viewpoint and that, combined with the perception of how surgeons behave, has led to a belief that surgical trials are not done properly. I think it is necessary to redress the balance. There are legitimate reasons why surgical trials are difficult to design and perform as a result of numerous pressures and problems. Martin Buxton has discussed several factors limiting our ability to perform RCTs in surgery but I would like to add several further factors to that list (Table 3.2).

Patient power is important. In the minimally invasive surgery case, patients heard through the media that it would mean a smaller inci-

Table 3.2 **Pressures and problems in performing trials**

- Patient power
- Industry leverage
- Third parties
- Funding
- Timing and delays
- Uncontrolled diffusion
- Personality issues

sion and that they could go home earlier. So it was difficult to find patients willing to accept the traditional treatment for a controlled trial. Patients were just not interested in whether they might have to have a further procedure later in the year if they had minimally invasive surgery, or whether it might be less effective.

There is a huge amount of pressure from industry to introduce new surgical technologies, particularly with new disposable devices. I also have concerns that the Medical Devices Agency (MDA) does not look at the quality of trials properly. A new piece of equipment comes to this country and it seems to be sufficient for some practitioners to say they are doing a trial on it in order for it to be produced.

A further pressure, that perhaps occurs in the US rather more than in the UK, is that surgeons are being pushed to use a technology by third parties. I know of several surgeons who have been told to use minimally invasive surgery because, if they did not, cardiologists would stop referring patients to their hospital and the hospital would lose revenue. These third parties are telling surgeons to adopt a technology without any evidence of its value.

Lack of funding for surgical research is an inhibitor to doing good trials. It is probably true that grant giving bodies have preferences for molecular biological research. The molecular biologists have stolen the intellectual high ground. I try to tell people that I am a 'molecular surgeon' – I operate on all the molecules at once!

A further disincentive to doing surgical research, particularly surgery that involves a high capital cost for an implantable device, has been the reluctance of device manufacturers to take a role in funding the capital cost for these new pieces of surgical equipment. If a company is going to benefit from a higher share price, they should have, within the 'D' part of their R&D budget, some way of paying for the capital costs of their device whilst it is being evaluated. Within the pharmaceutical industry this is easier, in that the individual cost of the trial drug is small compared with the total investment, but this may not be the case for complex medical devices. Companies want rapid regulatory approval so that they can quickly go to the market. Not infrequently, a good trial is designed and set up but ultimately fails because the equipment is given marketing approval and then the company no longer wants to fund the trial.

The timing of trials is important and it is always easy to say, 'We can't do a trial yet because there will be a better piece of equipment next year.' A good example of this was the introduction of intra-coronary stents to dilate coronary arteries. There were always going to be better stents next year. So, when should a trial be conducted? I think you have to hit the moving target at a time when a good audit has been done and the technology (not just the surgery) has been worked out. Then an appropriate evaluation can be performed.

There are also delays which could be avoided, both in getting funding for the trial, and then in completing the trials. One of the difficulties and frustrations is that, unless the diffusion of the technology can be controlled during and outside the trial, eventually the trial is extended and by the time it reports it has become redundant. This is either because the technology has effectively already been adopted by loose diffusion, or because a trial has been performed and completed abroad.

The term 'personality issues' in Table 3.2 is a surrogate for individual surgical attitudes. Surgeons have got to learn that performing appropriate trials requires communication with each other as well as with other health care professionals. Surgeons also have to see other research, medical and non-medical personnel as equals in these

endeavours. There are often excuses from physicians or surgeons such as 'why do a trial to prove what we already know?'. This is clearly an attitude that has to be changed.

Case study: a trial of TMR as a treatment for angina

Transmyocardial revascularisation (TMR) is a way of trying to bring blood to an ischaemic myocardium using a laser hole directly punctured into the heart to try and prevent angina. This technology has already been promoted in the United States as being effective. In order to perform this operation, a large piece of costly equipment is required. How can an RCT of this technology be performed? How should the research be funded? How should patients be recruited?

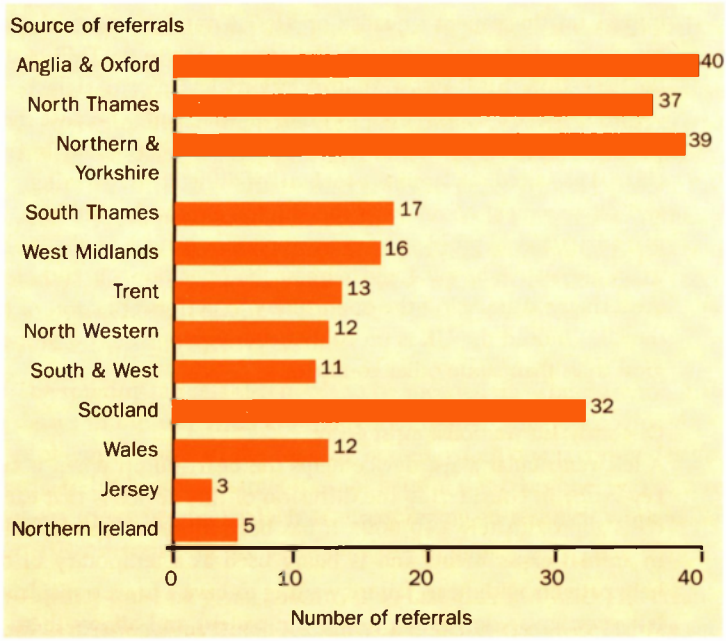
In setting up the randomised trial to compare TMR against medical therapy, necessary partnerships were formed with the Department of Health, other surgical centres, academics and industry. Different types of costs were identified for the trial: research costs, treatment costs and technology costs. The MRC funded the research costs, the NHS funded the service and treatment costs and BUPA bought the equipment. Without funding for the equipment, the trial would not have been possible.

The trial recruited patients from one centre but the recruitment was slow, taking several years. This suggests that in order to recruit the required number of patients expeditiously, multi-centre trials should be considered. One explanation for the slow recruitment in this trial can be seen by examining referral patterns from regions around the country (Figure 3.2).

The number of referrals can be seen to vary by region. In the area local to Papworth, there was quite a large referral base. But in other areas where there is more heart disease, fewer patients were recruited, with the exception of one excellent centre in Scotland. Clearly, if other centres in the UK had been involved in this trial the result would have been produced quicker and would have been more valuable.

Preventing the diffusion of unevaluated technologies is, perhaps, one of the most difficult issues to be addressed within surgical trials. One of the ways to control this is, of course, to advise NHS Trusts who have invested in these technologies to use them only in the context of a

Figure 3.2 **TMR trial – referrals by region**



trial. Such advice is, however, not mandatory and perhaps Trusts should be encouraged to heed that advice until the final trial results are available, either through incentives or perhaps through regulation. The question then arises of whether Trusts should be prevented from purchasing expensive pieces of equipment before the trials are complete. In the case of the TMR trial, several NHS Trusts went ahead and purchased the highly expensive equipment before the trial was finished.

Case study: lung volume reduction surgery

Another example of a well designed trial which was difficult to introduce, is that for lung volume reduction surgery. This is an operation to reduce the size of the lungs in patients who have emphysema, to try and

improve their physical well being for what is, as yet, an undefined period of time. If it were found to be successful, it could have a significant impact on the patient's quality of life. A randomised trial was put together in this country using a similar approach to the TMR trial, with the exception that it was a multi-centre trial. Unfortunately, however, the MRC made a decision not to fund the trial, partly because they had heard of a larger US trial that would probably report earlier. In fact, the US trial has faced the familiar problems of difficulty in selecting patients and disagreement about what the surgical protocol should be. There was an additional problem of crossovers since patients in the US can go to an alternative hospital and request the operation. It is therefore a great shame that we lost the opportunity to perform this sort of trial in the UK. Indeed the UK is probably better equipped to do proper surgical trials than some other countries.

Case study: left ventricular assist device

A left ventricular assist device helps the heart pump when it fails. A key factor in considering the diffusion of this device is that each one costs approximately £40,000. It has been found to be very effective in animal experiments and is being used as a temporary bridge to help patients with heart failure waiting to have a heart transplant. The left ventricular device sits inside the patient and allows them to be mobile. We tried for many years to set up a trial, which was designed for a small number of patients and was discussed with various government agencies. The aim of the RCT was to assess the benefits, costs and technical performance of the device as a permanent solution for patients with heart failure who are unsuitable for cardiac transplantation. The benefit measures included survival and measures of mobility and quality of life. The trial never got off the ground because nobody would fund the devices. The company was not interested in funding the device in the trial as it was already being marketed and sold around the world. The NHS, quite rightly, did not feel that it could invest in devices which were still in development, without input from the company. This conflict identifies a need for a better understanding and earlier symbiotic discussion between health departments, and device manufacturers so that proper evaluation can be built into the development of the device.

Models for change

Having examined some examples of good trial design, some of which have worked and some not, I now want to propose several models for change. Firstly, a multi-disciplinary approach to surgical trials is required, involving academic partners and health economists. Secondly, multi-centre trials and collaboration between surgical units is needed, since no single surgical unit is likely to see enough patients to recruit for a large trial without undue delay. Thirdly, proper R&D skills and support need to be promoted within NHS Trusts. Although this may seem expensive, there is no point in us continuing just to do today's technology well, we have got to learn about tomorrow's technology. Finally, teaching R&D skills to all health care workers, but particularly surgeons, is essential.

The funding process also needs to be speeded up. There are too many delays in the way trials are funded. Although this is recognised as a problem by some of the funding bodies, when one is trying design a trial to hit a moving target it is not helpful if it takes one, two or three years to set up the trial. It is discouraging to research workers and inevitably results in the trial being out of date.

Industry must be given incentives to take part in trials. The creation of the National Institute for Clinical Excellence may well provide such an incentive. We shall have to wait and see. There is of course the risk that industry may choose only to launch its technology in countries where it may be introduced without the requirement for evaluation, particularly the potentially bigger markets of the rest of the EU and the US.

If some of the above proposals were in place, then it would perhaps start to answer some of the questions of timing and funding and would bring together the partnerships of government, surgical centres, academics and industry that are so important.

NHS Trusts could be encouraged to further improve their role in surgical research. They have to demonstrate leadership and co-operation as well as a willingness to invest. They have to accept that there is an uncertainty of outcome. I think that Trusts have to accept that any new technology may not work, and that there is just as much value in

doing a trial to show that something does not work, as to show that it does. Thus there has to be progress in willingness to accept uncertainty. Proper management structures and support services are required. Perhaps not all hospitals and Trusts should be doing R&D, or at least not trying to lead it.

Better relationships with industrial partners are needed such that industry is involved in research but does not take over trials and insist that results are reported earlier. Further action is needed to speed up the health technology assessment process. Companies should be required to subject their innovations to a higher level of scrutiny. A dialogue with the MDA and with NICE could perhaps help with this.

Ways must be found to protect the public more effectively from unanticipated adverse effects of health technologies, including surgery. Whilst some progress has been made to achieve this and there are mechanisms in place to make further progress, the onus is on the surgical research community to promote better surgical research.

REFERENCE

Horton R (1996). Surgical research or comic opera: questions but few answers. *The Lancet*; 347:984-985.

Discussion

BRENDAN DEVLIN

In considering John Wallwork's paper, the first issue that came to my mind is the nature of evidence based surgery. RCTs are not the only form of evidence available for surgery. Other types of evidence exist, such as case controlled studies and prospective cohort studies. I think that a new rubric is needed for case controlled studies, and for case series cohort studies: how we assimilate the data in them, how much value we place on the data in them, and how they are used. Surgery involves not only the surgical intervention but also the decision to operate. Very good case controlled studies with good decisions to operate must be of some value. People do not go to a surgeon just to be randomised to have something done to them. The paradigm of a good surgeon is someone who can make a correct diagnosis, perform the appropriate operation and manage the patient well post-operatively. In promoting case controlled studies, collaboration with scientific methodologists is required in order to determine what surgeons should put into case controlled studies.

The European Surgical Research Society meeting in Hamburg in November 1997 spent a day discussing what constitutes evidence in surgery and no conclusion was reached. There is no conclusion to it. The end result, the fact of being a practising surgeon, is that surgeons deal with patients, and patients come along to surgeons because they trust the surgeon to do something. Patients do not want just to be randomised.

The tidal wave of technology and the peer pressures that go with it have been discussed by John Wallwork. How can this peer pressure be best handled? Has anybody thought about giving surgeons some support in handling this information about the effectiveness of surgical technologies? How do surgeons learn what is new? There is the question of specialty meetings, which John Wallwork mentioned, but how do they help? The whole problem is that we must keep up with peers and the specialty, both of which are often trying to do new things all the time.

Some people have suggested the way forward might be a network of guidelines; but guidelines have enormous problems. How is the evidence for guidelines graded? What evidence should go into guidelines, and what should be done with the guidelines once they have been produced? 'Private Eye' magazine recently pointed out that surgeon number 196 in the Royal College of Surgeons' recent audit of prostatectomy outcomes killed four times more patients than anybody else. What should be done about this? Should the Royal College of Surgeons go ahead and say, 'In future surgeon number 196 does not operate'? Or should surgeon 196 be retrained? Or is the question more about how we should be controlling for case-mix: surgeon number 196 may have treated the sickest patients?

Last year the Royal College of Surgeons examined cleft lips and palates surgery, and we identified quite clearly five locations where patients should not have cleft lip and palate operations done at all. How should this situation be managed?

Lastly, there is the whole problem of the learning curve. At what stage of the learning curve should a trial be performed? Surgeons' technique also changes and improves throughout their careers. Should another randomised trial be performed later? I have no solutions to these issues. They will be with us for some time yet.

Chapter 4
How should the NHS obtain and exchange information on emerging surgical techniques?

PROFESSOR JOHN GABBAY

Introduction

In recent years there has been a big shift of emphasis from the idea that all cost-effective treatments should be free and universally available, to the idea that all treatments that are free and universally available ought to be cost-effective. Another little irony, perhaps, is that there is a tendency to talk about getting research into practice, but actually what should drive a useful research programme is to get practice into research. The objective of R&D is to answer questions that are of practical interest to those on the front line in the NHS. The aim of this paper is to describe how different parts of the NHS have been trying to do this.

The organisation of research: national, regional or local?

There are three potential levels at which a research programme in health technology can be organised: national, regional and local. There are pros and cons to research at each level, depending upon the research aims. If the aim is good research involving high quality multi-centre trials and large recruitment, then clearly the local level is not appropriate. If the aim is to perform a systematic review, individual local areas or regions are not suitable since these types of research should be performed at a national level, assuring the quality of the research and avoiding duplication. The potential drawback of research at a national level, however, is that it may not be as responsive to the real questions faced by the practitioners who need those questions answering. National research may not address the right topics or not address those topics at the required speed. There is also a question of how acceptable national results are to people working in local settings with different facilities and resources.

In practice, some research needs to be conducted at each of the three levels: national, regional and local. I will discuss work at the first two of those.

Table 4.1 **The NCCHTA**

- Identifies HTA questions
- Supports their prioritisation
- Commissions research
- Monitors the research
- Disseminates the products

Research at the national level

The National Co-ordinating Centre for Health Technology Assessment (NCCHTA) supports the Standing Group on Health Technology in deciding on the NHS's research priorities. The NCCHTA has five main tasks as listed in Table 4.1.

It tries to identify the appropriate questions in health technology assessment and to support the work of prioritising them. Approximately 1,400 ideas for the health technologies that ought to be assessed are received by the NCCHTA each year, and they need to be reduced to a manageable number. There is then a commissioning process which is often very long, arduous and complex. Commissioned research bids are monitored to ensure that the research is being done adequately and that the methods, which are often very developmental, are working, and to keep the research projects on track. Then, when the results of the research are ready, they have to be disseminated and implemented. I will discuss each of these tasks in turn.

Identifying HTA questions

The first task is to identify the possible research topics. These have to be questions that matter to the people in the NHS and where there are real evidence gaps. In identifying the topics, there is a complex procedure of very widespread consultation (in fact, some would argue too widespread) and some much more focused consultation. For example, focus groups of general medical practitioners (GPs) are used, which reveal some useful ideas for HTA. Documentary sources

are also used and a procedure called horizon scanning, which I will discuss below.

This part of the process gives some idea of where the evidence gaps that really matter are. There is then a very rigorous programme of trying to refine the questions, clarify them, and set clear priorities for which are the most important. Committees or 'panels' of experts, arranged around broad areas of health technology such as acute care, are involved in this. Topics that are identified through this process are usually problems that have already hit the NHS. The Annual Report of the HTA programme lists all the topics that have been prioritised.

Horizon scanning

Rather than identify topics that have already hit the NHS, it makes a lot of sense to scan the horizon, to look further ahead and see what is likely to be impacting on the NHS in future. Then there is time to do the research and have the answers, so that we do not end up with a similar problem to that which has arisen, for example, around beta interferon, which was suddenly upon the NHS before a policy on its use was ready. The idea of horizon scanning is to provide an early warning system on new and emerging technologies which are likely to affect the NHS in five years' time and beyond. Several initial methods of horizon scanning have been adopted and are discussed in Stevens *et al.* (1997) and Robert *et al.* (1998).

The first step was to search journals and conference abstracts. The results and experiences of other countries with horizon scanning were also examined. Second, a watching brief was kept on the development and diffusion of new pharmaceuticals. It is relatively easy to do this for pharmaceuticals since there are various hurdles they have to go through before they can be registered, such as Phase I, II, and III trials. These alert us to the evidence and potential of new medicines, but that is not enough. We also conducted a postal survey of some 3,000 individuals in the UK, asking in effect: 'What technologies do you think are going to be important? When do you think they are going to be important? What is likely to be the impact – major, moderate or small? What are the reasons for this? How well evaluated do you think it is?'. Figure 4.1 gives an example. Although

Figure 4.1 **Questionnaire**

New/ expanding technology (device, drug, procedure or setting)	Will be of importance a: now b: 1996-7 c: 1998-2000 d: 2001	Impact for NHS a: major b: moderate c: some	Reason for its impact a: benefit b: total cost c: organisational d: rapid diffusion e: other (please specify)	How well evaluated a: fully (convincing RCTs) b: quite well c: partially (some comparative evidence) d: not well e: don't know	Please name one other expert on this topic (name and place of work)
Please name the technology	**Please ring the most appropriate letter**				
Example: Beta interferon for multiple sclerosis	a (b) c d	(a) b c	a (b) c (d) e	a (b) c d e	

these questions were very broad, they did give some idea of the new and emerging and technologies.

Horizon scanning was conducted in 1995 and 1997 and further horizon scan is ongoing. The responses for surgery for 1995 are shown in Table 4.2 as an example. It seems to me not a bad list for predictions made a few years ago.

Several methodological issues arise in horizon scanning. There may be an over-representation of certain specialties in the survey sample. The questionnaire was sent out using a cascade method where medical directors and NHS Trust chief executives were asked to distribute it to the people they thought most appropriate, and this led to certain specialties, such as surgeons, physicians and radiologists, being prominent. There is also the problem of how people interpret the word 'technology'. There tends to be an emphasis in people's minds

Table 4.2 **Results of 1995 horizon scan (surgery related topics)**

Minimally invasive surgery (2)*
Implantable vascular stents (4)
Doppler measurement studies (7)
Laser treatment for benign prostatic hyperplasia (8)
Telemedicine (11)
Interventional radiology (13)
Angioplasty (14)
Lasers for dermatology (16)

Note: *Overall ranking

on technology as some kind of technical apparatus or gadget. Of course, health technology can mean devices, drugs, surgical procedures, or how services are organised. Another issue with horizon scanning is that most people's horizons are fairly close and respondents tend not to look far enough ahead. They may also neglect less well defined technological areas.

A further methodological issue is the reliability of the sources. Good sources for horizon scanning have been explored at the Wessex Institute for Health Research and Development (see Robert *et al.*,

Table 4.3 **Good sources for scanning the surgical horizon?**

- Specialist medical journals
- Principal medical journals
- Medical engineering companies
- Devices companies
- Private health care providers
- Newsletters and bulletins from other national and regional HTA agencies
- Sentinel groups of expert health professionals

1999). Sources for surgery are summarised in Table 4.3 and include journals, devices companies, private health care providers and newsletters from other similar exercises in other countries.

Dissemination

The NHS Centre for Reviews and Dissemination (CRD) at the University of York performs some excellent research on behalf of the NCCHTA. The CRD also produces the Effective Health Care Bulletins. It usually takes a couple of years to produce these bulletins, given the rigorous methods required to do thorough, systematic reviews. Yet health authorities and NHS trusts need the answers, preferably last week and certainly next week. This timing issue needs to be addressed but it is clearly going to be difficult to resolve.

Research at the regional level

Development and evaluation committee

In the former Wessex Region of the NHS a few years ago, a scheme called the Development and Evaluation Committee (DEC) was set up. Other NHS regions then took up the idea. The work of the DECs has been taken over by NICE since 1999. The DECs themselves have been shut down and their scientific teams now provide reports to NICE. Although the organisations have changed, the work remains the same.

The DEC was an attempt to help purchasers and providers decide which new technologies were or were not worth investing in. There were three key elements in the process. Firstly, clinicians and health care commissioners were involved in choosing the topics of importance and relevance to them. Secondly, a scientific team at the Wessex Institute (for example) produced a review of each selected topic within three months. The Wessex DEC produced 50 or more reports over the few years up to and including 1997. The results of the DEC reports have been compared with the results of later systematic reviews, and so far there have not been any different answers. Thirdly, a regional committee comprising influential clinicians and managers pronounced upon the technologies in the light of the DEC reports. Committee members were distant enough from the local politics of the individual Trusts, where clinicians may want to want to push their

own ideas and Trust managers to conserve resources, but close enough to know what the regional issues are. The committee members were sufficiently local and well-respected that the local health care commissioners and clinicians could and did keep in touch easily.

Each of the Wessex DEC reports consisted of a critical review of the evidence, followed by a statement of the quality of that evidence. Having assessed the evidence, the DEC report came up with an estimate of cost utility, in terms of a cost per quality adjusted life year (QALY). Despite the known limitations of the QALY approach, its does least give some indication of value for money.

The strength of the evidence was classified according to whether, for example, a RCT had been conducted (Table 4.4).

The DEC reports use five categories of recommendation: strongly recommended, recommended; borderline (beneficial but high cost); not recommended; and not proven. A matrix was used to classify technologies into these categories where the columns represent the cost per QALY figure and the rows represent the quality of the evidence. The interaction of the rows and columns gives the category of recommendation (Figure 4.2).

The DEC approach has been applied to surgical techniques. In 1993, out of 19 DEC topics, four were surgical: lithotripsy for urinary stones was recommended; adult cochlear implants were strongly recommended; for prostatic stents no decision was made; and sterile off the

Table 4.4 **Quality of evidence categories**

I	At least one properly designed, randomised controlled trial
IIa	Well-designed controlled trials without randomization
IIb	Well-designed, cohort or case controlled, analytic studies, preferably from more than one centre or research group
IIc	Multiple time series, or from dramatic results in uncontrolled experiments
III	Opinions of respected authorities
IV	Evidence inadequate, owing to problems of methodology or conflicts of evidence

Figure 4.2 **Categories of DEC recommendations**

Evidence	<£3,000 per QALY	£3,000 - £20,000 per QALY	>£20,000 per QALY	Negative QALYs
I	Strongly recommended	Strongly recommended	Borderline	Not recommended
II	Strongly recommended	Recommended	Borderline	Not recommended
III	Recommended	Borderline	Borderline	Not recommended
IV	Not proven	Not proven	Not proven	Not proven

shelf bone grafts were recommended. In 1994, again out of 19 DEC reports, five were surgical topics. Decreased D&C (functional bleeding) was strongly recommended; grommets for glue ear was not proven; surgery versus elastic compression for varicose veins was not proven; microwave benign prostatic hyperplasia was not proven; screening for prostate cancer was not recommended. It is interesting to note that most surgical technologies end up being recommended or not proven. There are very few that are not recommended, although one such was prostatic screening – an example of the DEC getting the results right before the big studies came out. For 1995, there were four surgical topics out of 16; in 1996 two out of 17; and in 1997 three out of 25.

An examination of the breakdown of all DEC assessments for 1995/96 gives an idea of the kind of pattern of reports and recommendations by broad classifications of technologies. Of the 32 reports submitted in 1995/96, 16 examined drugs, seven procedures, three devices, four screening procedures and two concerned the settings in which health care is delivered. For the 32 reports, the most frequent recommendation category was 'not proven' (13/32), highlighting an important fact: that sufficient evidence is not there in almost half of the topics studied in detail. Seven of the 32 assessments were 'strongly recommended', five were 'recommended', six were 'borderline' and just one was 'not recommended'.

The key features of this system, which made it work very well, were that:

- there was a 'bottom up' supply of topics;
- the people at the front line informed the DEC what they needed to know about;
- the reports were produced within three months; and
- the committee pronounced within six months from starting the process.

Thus there was a rapid turn-around and health service practitioners knew that if they asked a question they were going to get it answered reasonably accurately and quickly, assuming it was answerable at all. A further important feature was the clarity of the recommendation.

Although our application for funding for a formal evaluation of the DEC was turned down, feedback from our 'customers', suggested that they, particularly health care commissioners, did use the DEC reports to guide decision making. Comparisons with the 'gold standard' reviews that are now coming from the CRD and NHS HTA programme do seem to suggest that most of the time the DEC answer was right. In future, it will be interesting to see how NICE's appraisals of new technologies compare with the various DECs' findings.

REFERENCES

Robert G, Gabbay J, Stevens A (1998). Which are the best information sources for identifying emerging health care technologies? An international Delphi survey. *International Journal of Technology Assessment in Health Care*; 1998 Fall; 14(4): 636-643.

Robert G, Stevens A, Gabbay J (1999). 'Early warning systems' for identifying new health care technologies. *Health Technology Assessment*; 3(13).

Stevens A, Robert G, Gabbay J (1997). Identifying new health care technologies in the United Kingdom. *International Journal of Technology Assessment in Health Care*; 1997 Winter; 13(1): 59-67.

Discussion – the role of SERNIP

SIMON WOOD

By way of an extension to John Gabbay's presentation of mechanisms for information exchange in the NHS, I want to discuss a development for information exchange that is specific to surgery: SERNIP (Safety and Efficacy Register of New Interventional Procedures). SERNIP is organised by the Academy of Medical Royal Colleges and has been functioning since 1996. In setting up SERNIP the basic but key questions of how to define safety, efficacy and new procedures had to be answered. It was clear that SERNIP's concern is not with effectiveness (how a new procedure will perform, given all the different spectrum of different abilities of surgeons and investigators). Nor is its concern with cost issues. SERNIP is upstream of that, looking at very new developments: whether they seem safe and whether there is good scientific evidence of efficacy. Where there is evidence on a new surgical technology, then SERNIP considers whether it should be helped downstream into the developmental process.

When SERNIP was first mooted, some in the surgical community felt that it was another tier of regulation, another hurdle to get over. The philosophy of SERNIP is quite the reverse, however. It aims to identify promising and safe new techniques, diffuse them quickly to a limited number of centres to perform co-ordinated studies in order to generate some good hard scientific data, and then to disseminate the findings.

When SERNIP is notified about a new procedure, we make a synthesis of the available literature on it. This report is then put to a committee of representatives from all relevant Medical Colleges who sit quarterly and assess whether, for example, 'It is looking good and safe and can be used' (category A), or whether 'There is not enough data, it needs more trials' (category C), or whether 'It should not be done' (category D).

One feature that distinguishes SERNIP from the NCCHTA's approach is that the answers are reached relatively quickly. The NCCHTA process (prioritisation, commissioning, monitoring, reporting and rec-

ommendations) is, of necessity, lengthy albeit very thorough. This can result in a new technology awaiting good scientific data which may take a long time to be collected, during which time the technology may have started to diffuse. One wonders what happens when, at the end of the three years, the results of the research are published and the conclusion is that the technique should not be used.

SERNIP has a different approach in that the new technique is identified early and the aim is then to see whether there is an inherent problem with it. If there is, SERNIP then aims to address that problem and its implications as quickly as possible.

Industry is heavily involved in the promotion of a new technology and dealing with industry forms a much larger part of SERNIP's work than originally anticipated. In the first six months of SERNIP's existence, 40% of inquiries came from industry. The medical device suppliers are involved in promoting and publishing research on new procedures but in many cases their agenda is more commercial than scientific and therefore SERNIP liases closely with the MDA.

The second most frequent source of inquiries is from commissioners of new technology (public health medicine departments, private medical insurers) who want to know whether they should pay for it or not. If, as result of the literature review, a procedure is in evident need of more research, then our recommendation is that it would be appropriate to commission the research only from a centre where we know that the necessary studies are in progress.

Does SERNIP work? I believe that it has had an impact. SERNIP is a professional, non-governmental organisation which is also voluntary and advisory. I think that probably the best example I can give is that of new thermal endometrial ablative techniques for menorrhagia. There were four different devices considered in need of further study, on both safety and efficacy grounds. They were all put in the 'C' category (not enough data and requiring more evidence). So what then happened? Firstly, the leading clinicians involved in all four techniques agreed to co-ordinate multi-centre observational studies. The device suppliers responded differently. One company decided to hold off marketing until they had a good RCT against a comparable treat-

ment. They have now produced and published the RCT results. A second company went ahead with several multi-centre, multi-national studies and are holding off marketing until the results are known. The third company felt that there was no chance of getting any data and have withdrawn the product from the UK. The fourth company was told that SERNIP was not happy with the safety standard of the device and was advised to modify the design. Within the space of about six months, the SERNIP process altered both the commercial and the academic progress of these devices and, I hope, facilitated their development towards safe and effective use rather than impeding their progress.

Chapter 5
How is surgical R&D delivered in practice? A Trust view

PROFESSOR AL AYNSLEY-GREEN

Introduction

The papers so far have discussed R&D at a fairly macro level, looking at the philosophical issues of surgical research, and at national and regional priorities. But in the end what actually matters in terms of delivering R&D is the culture in which the R&D takes place, the organisational procedures which support it, and the philosophies which underpin it. The structure of this paper is threefold. First, it sets out the issues that Great Ormond Street Hospital for Children (GOS) has had to face over recent years concerning how to deliver first-rate clinical R&D. Secondly, I extract from that the issues that are generalisable and applicable to the dilemmas in delivering surgical R&D. Finally, I conclude with a series of proposals identifying components of a possible way forward.

The institution

The Institute of Child Health (ICH), which is funded by the Higher Education Funding Council of England (HEFCE), and GOS, which is an NHS Trust and receives its support from the Department of Health, are joint partners. There is a determination to integrate the R&D activities in both the ICH and GOS and to be an indivisible institution. Although each has its own paymaster, the activities of both are directed at weaving together the clinical, research and teaching excellence of our colleagues. There may be messages from this for other centres in terms of how NHS Trusts relate to their local university. Although the ICH and GOS have the privilege of being close geographically, and are dealing with the common issue of child health, nonetheless the philosophy of integrating university activity with Trust activity is, I would argue, fundamental for success in delivering research of high scientific standing and value. There is a growing concern that the preoccupations of the Culyer report with Trust costs may threaten this integration.

Imperatives for change in the delivery of R&D

In 1992, there was increasing concern outside the, perceived to be rather arrogant, ivory tower of ICH/GOS about its role, its position, and the deliverables of child health R&D. There was increasing uncertainty about what the ICH and GOS stood for, about the quality of their R&D, and about what it meant for improving child health in the future.

The first blow to ICH/GOS was the 1992 HEFCE research assessment exercise (RAE) which rated the ICH's performance overall at 3.2. This was really quite a shock to an institution that thought itself to be very good and delivering good research.

A further blow was when, in the same year, the government faced the dilemma of what to do about the Special Health Authorities (SHAs) in London. It created the Thompson Committee, which assessed the productivity, value, quality and relevance of the research being performed in all eight SHAs. GOS received quite a savaging from the Committee over the quality and relevance of its research, with much being considered to be of low scientific credibility and of low relevance to the NHS. An important message to come out of that savaging was the fact that the evidence presented to the Committee had largely been by system specialty, for example cardiology or general surgery, rather than by common research themes. The importance of the presentation of the research programme is a lesson that has been learned. A further reorganisation was the change from being a protected SHA for the best part of 26 years to suddenly becoming a NHS Trust and being thrown into the competitive NHS internal market. Finally, there was the spectre of Culyer emerging on the horizon and the importance of central support for R&D to NHS trusts.

What emerged from 1992 was a series of imperatives. We had to improve the scientific quality of basic science and clinical R&D. Then we had to improve their relevance to the NHS. Thirdly, we had to demonstrate the value of the R&D, and so ultimately to justify why GOS and the ICH should continue to survive.

Changes to the organisation and delivery of R&D

Central to the way forward was a recognition that there needed to be the closest collaboration between the Dean of the ICH (Professor Roland Levinsky) and myself as the new Director of Clinical R&D. The new philosophy of integration was exemplified by portraying ourselves as an indivisible partnership, in the pursuit of joining together basic and clinical scientific strength.

There were four major tasks to be completed. First, to organise R&D. Five years ago the institution did not know precisely what research was going on, other than the grants that were channelled primarily through the ICH, and so we had to develop processes and define the research portfolio. Second, we then had to verify the existence and the quality of projects. Third, we had to deliver a research strategy for the future, to justify our level of support. Finally, we had to present our performance both internally and externally – a vital task and one that is often neglected.

Philosophy and objectives of the institution

Before addressing the details of the organisation of R&D, the institution had to revisit its underlying philosophy and objectives. The mission of the ICH and GOS is very simple: to improve the health of children. But what about the principle underlying this?

I would like to introduce you to what I call the 'Alexandrian Philosophy'. A full-page advertisement which appeared in the Sunday Times during 1997 said: 'Think big! Whoever heard of 'Alexander the Average'?' The importance of the sentiment is of course that it is 'Great' Ormond Street Hospital not 'Average' Ormond Street Hospital! An organisation needs to have a vision and an expectation of quality and excellence. Our, fairly modest, overarching objective is to be the leading centre in Europe for R&D in paediatric medicine and child health.

In our research we are trying to understand what determines healthy development in normal children; to understand the processes which lead to disease; from that to define new and better ways to diagnose diseases, especially without causing pain, an aspect which is very

important for children; and finally to discover new ways to treat children to improve outcome and quality of life. These research objectives have been agreed by colleagues across the institution.

Organisation of R&D

The indivisible joint partnership of the ICH and GOS means that all staff are equal partners in R&D. This is a very important concept because in the early 1990s the only research that was thought to be of value was that arising from high technology molecular biology. Colleagues in clinical disciplines felt disadvantaged. Everyone has to have an equal stakeholding in the R&D profile, with a partnership of skills. Clinical excellence is essential since without it patients will not be available for research. At the same time, we need other colleagues with the training and experience to be able to raise prestigious programme grants from the MRC and other funding bodies. The conundrum every institution and department faces is how to marry these skills together so as to push forward a portfolio of high quality R&D on a broad front. What about encouraging research in nursing and the professions allied to medicine? These professions are very undervalued, hugely important and yet under-resourced.

Organisational professionalism is essential to successful R&D. This means having a clear objective and a strategy for reaching it. Central to this is to weave together epidemiology, basic and clinical science, and health evaluation.

Turning now to the practicalities of R&D, a big headache is determining the requirements for first-rate R&D. What is 'R'? The popular view of 'R' is the boffin in the laboratory with the high tech medicine. But the Oxford English Dictionary defines 'research' as 'systematic investigation to establish facts or collect information'. If this is the case, then the whole organisation has to sign up to the view 'we challenge what we do'. That is what research is all about. Much of research is inevitably implicit. It cannot attract external funding. Herein lies a central dilemma with the Culyer process, which is, how can implicit research, which may be very valuable to the NHS, be conducted without receiving external funding?

What about 'D'? The Department of Health provides no help in terms of defining 'D', so the definition proposed at GOS/ICH is the 'systematic evaluation of the application of the results of research into practice'. This is different from audit. This definition of 'D' is important surgically because there much of the work that goes on in the institution is very effective 'D' rather than 'R'. Recognition of this concept needs to be achieved in order to ensure appropriate funding for its delivery. At present colleagues are being asked to register both types of project.

Figure 5.1 **What are the requirements for R&D?**

- Culture
- Organisation
- Principles
- Resources
- People

Figure 5.1 lists the five key components of the organisation for R&D: its culture; the organisation of its R&D; the principles which underpin its activities; the resources it needs for its R&D; and last, of crucial importance, the people who actually work in the organisation.

In terms of culture, there need to be levers to pull, and the levers in this instance were financial and the threat to survival from the shocks given to the institution in 1992. Without these, the cultural change since initiated could not have been introduced. The second prerequisite is a commitment from the highest of levels in the organisation that R&D actually matters. The way of demonstrating this in our organisation was to create an executive directorship on the Trust Board for the Director of Clinical R&D.

Then there has to be change, and support for change, throughout the organisation. People who thought that they were very good, suddenly found after external review that they were not considered to be so. Immediately peer pressure, the insult to self-esteem and the incentive

for excellence, created a reaction that was a great driving force for support behind change.

Money to invest in people is fundamental, and this bedevils most of the regional undergraduate teaching centres which had funding cut after the HEFCE research assessment exercise in 1992. It is very difficult now to find new money with which to invest in the future. GOS/ICH has advantages in this respect: the emotive appeal of sick children and access to a very powerful fund-raising engine. As a consequence, we have the potential to raise large amounts of research money. Fund-raising relies on effective marketing, which in turn requires portraying a vision for the future. Many specialties have been remarkably backward in thinking about the marketing analogy of how to raise money from a variety of sources and not depend solely upon central government.

Organisationally, we have a Director of Clinical R&D and a Dean, who work together, and a Joint Research and Strategy Committee which serves both of the partners in the institution. At the specialty level, there is a lead clinician who carries the responsibility for interfacing with the research and strategy committee, and academic unit heads to whom all member of clinical and non-clinical staff report. A mandate was issued by the Board of Governors of GOS and the ICH's Committee of Management that R&D was important and had a high priority for development. The sanction levied was that only research registered through the newly established R&D office and fulfilling its R&D registration process would be supported in the event of any medico-legal problem in the future.

The rewards were difficult to sell to the staff to begin with. There was certainly the incentive of peer pressure and professional excellence, but of course what matters most of all is money. This was addressed by arguing that with the system about to be implemented more money could be delivered to the directorates, particularly through the commercial exploitation of intellectual property. In order to facilitate this, a full-time commercial administrator and negotiator was appointed, responsible for organising the portfolio of commercial projects, beginning with the financial contract and then monitoring

performance until the end of the project, monitoring the recruitment of patients into trials, etc. In general, industry has welcomed this professionalism, and our commercial portfolio has increased substantially over the last few years.

The process of R&D administration begins with data collection. An office was set up with a Director of Research Administration, with the appropriate staff and other resources, to collect information through a network of research co-ordinators in each of the specialities and groups across GOS/ICH. We have produced, in a common house style, a number of booklets and publications so that nobody can be under any illusion about what is expected when undertaking R&D at GOS/ICH. For example, every member of staff has access to a document defining exactly what has to be done to register a research project. It involves every member of staff reporting to an academic unit head who then peer reviews and physically signs off the quality of the project to be registered, in terms of protocol, hypothesis, statistics, ethical approval, etc. To facilitate good study design, we also provide brochures to advise on, for example, how to calculate sample size and how to develop appropriate questionnaires.

Verification of research activity and its quality

In the early 1990s the institution did not know what research activity was going on. In order to put this right, colleagues were first asked to inform the institution's R&D office of the titles of their research projects: 1,600 were identified. As this information flooded in, we realised there was probably some wishful thinking going on. So the existence and quality of these 1,600 projects needed to be verified. This was accomplished in a variety of ways. The first was a 10% random (well, not entirely random!) audit of the titles, and investigators who had registered projects were invited to meet with the R&D office to discuss their research protocols. At this meeting investigators were also asked to demonstrate that their research had ethical committee approval; that data was being collected; and that their research was being productive in the sense of being presented and published. As a result of this process, the research portfolio was refined and reduced to 1,100 registered projects.

The next exercise engaged in to verify research activity was patient tracking to identify the proportion of GOS patients involved in research protocols. This was achieved by accompanying clinicians on their ward rounds and entering the names of patients against research projects. In some specialties, such as oncology, 95% of the patients were engaged in one or more research protocols. The lowest proportion was 30%. Overall, in 1995, 85% of our patients were enrolled in research. This process allowed us to propose ways of making best research use of our patient base without exploiting the children. We found that most children were engaged in one or two projects, but some were engaged in up to nine. Number does not mean necessarily intensity or risk, but this approach provides a means to define what is actually going on. The disadvantage of this process is that it is a very labour intensive exercise. Ways of making it easier to access such data by incorporating it into our Patient Information Management System (PIMS) are being explored.

Finally, in terms of verifying activity, further information on research activity is obtained through the R&D costing process which is necessary for the submission of the bid for central R&D support funds in the Culyer process.

Developing a research strategy

There are five aspects to the R&D spectrum. First, the epidemiology of the condition: why is it important to study it; what is the burden of the disease? Second, what is the basic science in terms of understanding fundamental processes and pathogenesis and new approaches to diagnosis? Third, what is the clinical science of management? Fourth, what about health evaluation in respect to the impact of research findings? Finally, how can research be put into practice? The aim is to interweave each of these five dimensions into the research strategy and into the presentation of the research.

An organisation needs to stand back and identify the key questions that affect the health of the population it relates to. This was the approach adopted at the ICH and GOS, and a list of the important questions affecting the health of children was produced. Functional groups were built around these, bringing together colleagues from

Figure 5.2 **Seven interdisciplinary themes**

- Cancer
- Nutrition
- Inherited diseases and congenital malformation
- Infection and immunity
- Cardiovascular and respiratory sciences
- Neurosciences
- Population health sciences

different disciplines. This is a very different approach to the system specialty based approach that proved so disastrous in the 1992 assessment by the Thompson Committee. There are now 126 registered functional groups of varying levels of sophistication and of tenure: some are short-lived, others are much longer-lived. The functional groups are aggregated into seven research themes (Figure 5.2). They are broad themes, of course, but they reflect the major health issues of children. In each of these themes surgeons play a very important and prominent part. We are trying to get the hospital's delivery of services to be coterminous and intertwined with the research activities within the seven themes.

R&D is relevant to all members staff and groups of people, from undergraduates right the way through to purchasers. The medical staff issue is a very important one, and there are serious problems in generating an effective medical R&D workforce. There are problems with recruitment; the abuse of the lecturer post; difficulty in retaining staff; the paucity of rewards for R&D; and, of course, problems with succession planning. The issue of succession planning was drawn forcefully to our attention two years ago when new money was found for two new senior lecturer posts in important paediatric subspecialties. We advertised world-wide but no appointment was made because the applicants did not carry the portfolio expected of senior lecturers. The extent of the problem of succession became more

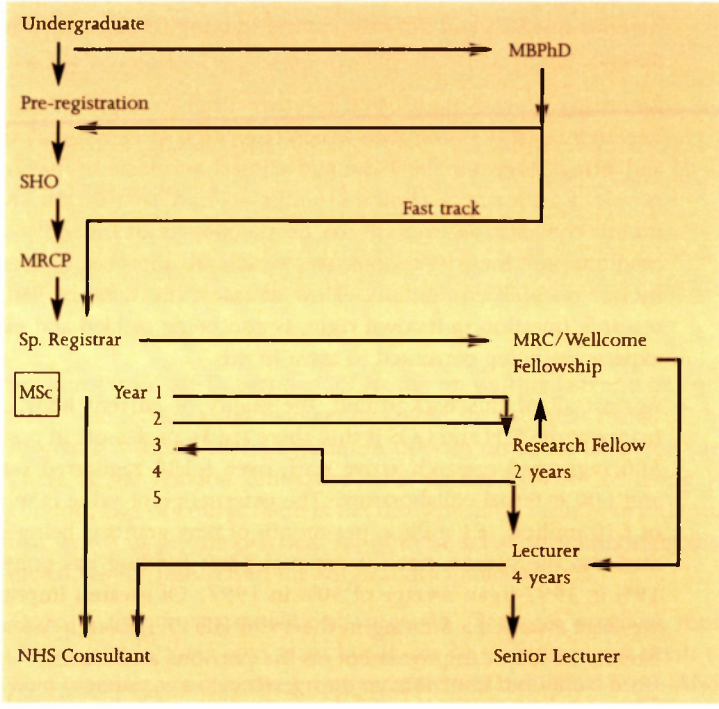
apparent when, prompted by this failure, I identified all of the senior lecturers, readers and professors in paediatrics in the North Thames region and found that of the 100 people with these titles, most would retire in the next five to ten years. At the same time there were only 20 lecturers in North Thames, none of whom was securely funded. Only six in GOS had proper protected time for research training as well as clinical training; most were soft funded; and most were 'jobbing' Registrars who carried a heavy teaching burden and were on rotations with other NHS Registrars. So where would the people come from to replace the retirees?

Using this information in child health, the postgraduate dean in North Thames was approached by senior professors in all seven institutions in North Thames arguing the case for the lecturer grade. As a result, we agreed on the concept for paediatric training shown in Figure 5.3.

That all is not well with research training in child health and paediatrics has been confirmed in a paper in the Archives of Disease in Childhood in December 1997. This sets out a survey of the experiences of paediatricians who obtained an MD in the UK over the preceding ten years. The MD has come to be regarded as an entry point for a consultant post. The analysis of the survey makes dismal reading in terms of: the quality of supervision; the time taken for the MD; and the quality of the outcomes of the MDs. Many doctors evidently experience second-rate supervision, second-rate research and second-rate outcomes.

There are two groups of people who need research experience. The first group are the 'ordinary' NHS consultants, who do not need to spend two or three years in research post but could have their research experience delivered by enrolling in an MSc programme during years one and two of their specialist training period. At GOS we have created just such a two-year MSc programme in paediatric clinical sciences, which includes teaching on the research process, its principles, and the statistical and ethical methods which underpin it. In addition, each student produces a research project. Every Registrar who rotates through GOS is offered the opportunity to enrol on the

Figure 5.3 **Training programme**



course. This intense, carefully audited programme equips most NHS consultants of the future with a lifetime understanding of the principles of evidence based medicine.

The second group are people who want to have a more intense research training by holding a research fellowship for at least two years. (I personally have no time for the one-year research experience). Following this, as a post-doctoral clinician, it is then possible to become a lecturer. Having been supervised, there are then four precious years to develop ideas and become a credible applicant for a senior lecturer post in due course.

In addition to agreeing these training principles, eight new paediatric lectureships have been created in North Thames, with 50% protected time for research and 50% for clinical training for accreditation purposes.

Returning now to the GOS/ICH's own R&D strategy: we created a Research Strategy Committee which covers the seven research themes and brings together the basic and clinical scientists by having two people, a basic and a clinical scientist, to lead each theme. An 18-month consultation exercise has been conducted to identify short-, medium- and long-term objectives which are approved and 'owned' by our research community. How to assess the value of R&D, as a research question in its own right, is also being tackled and external experts are being consulted to assist in this.

Against all of this background, the reality of current R&D performance at the ICH and GOS is this: there are 34 professors in post; over 500 registered research active staff; over 1,100 registered projects; and 600 external collaborators. The external grant value is in excess of £30 million. £1 million per month of new grants is being generated and the success rate for external grant funding has gone from 19% in 1992 to an average of 50% in 1997. Of greatest importance, we were awarded a 5 rating in the 1996 HEFCE Research Assessment Exercise, a major improvement on the previous assessment.

Thus, I believe that it is justifiable to say that, with the principles and policies I have described, and with the appropriate focus and planning, as well as with organisational commitment, it is possible to transform a poorly rated institution into a vibrant and important R&D resource.

Lessons for surgical R&D

Several lessons have been learned at GOS/ICH which are generalisable and applicable to surgical research. All surgeons at GOS/ICH are valued contributors to the institution but there are specific problems associated with them, especially in the clinical context (Figure 5.4). The clinical pressures are enormous: the limited availability of time for research; the conflict of service versus teaching versus research;

Figure 5.4 **Problem 1 – The clinical context**

- Clinical pressure – available time, the perspective of timescale
- Service vs teaching vs research
- Department of Health pressure – waiting lists, contracts
- Technical skills imperative
- Multiple site activity
- The London dimension
- Private practice

pressures from the Department of Health on waiting lists and on contracts; the technical skills imperative; and the fact that many of the surgeons working at the institution operate on more than one site. There is the London dimension, that is, the expense of living in London; and of course there is the distraction (if I can hesitantly use that word) of private practice. All of these act as very important disincentives for research in the surgical disciplines at GOS.

There is also the personality of surgeons. They are excellent team players within their own team but it can be difficult to get them to weld together as a cohesive group outside their individual team. Also,

Figure 5.5 **Problem 2 – Surgical culture**

- Personality
- Team values
- Rewards and stature
- 'Hands on' vs laboratory expertise
- Critical mass
- Research, development or audit?
- Declaration of R&D

Figure 5.6 **Problem 3 – Education and training**

- Technical skills focus
- Requirement for research exposure
- Opportunity for 'proper' research exposure
- 'Value' of research experience
- Abuse of lectureships
- Paucity of training fellowships
- Quality of research supervision
- Stature of academic surgery

many surgeons gain their rewards from peer approval of technical skill, or the monetary benefits of private practice, and so are less attracted to research. There is also the tension between time to maintain hands-on cutting skills and that for research. Other issues arising from the surgical culture are listed in Figure 5.5.

Other problems relate to surgical education and training for R&D (Figure 5.6). What is the Royal College of Surgeons doing to encourage a research culture in its trainees? As I have already discussed, I believe that we need explicit recognition of the importance of time for proper R&D training. This should be a criterion for the accreditation of clinical training programmes, including those for surgeons. We also need more lectureships and training fellowships in surgery.

The nature of surgical research is broader than merely introducing innovative techniques or prostheses. It extends to management of disease outside the operating theatre, and to the other aspects listed in Figure 5.7. These are all important and we are trying to weld all of them together in our R&D work at GOS/ICH.

Conclusion

In conclusion, I propose some possible solutions to the malaise described above. First, a cultural re-alignment on the importance and

Figure 5.7 **Problem 4 – What is 'surgical' research?**

- Surgical techniques – evolution vs revolution
- Management of disease states outside the operating theatre
- Delivery of health care
- Basic science of surgical circumstances and diseases
- Evaluation of the application of 'R' to 'D'

value of research in clinical surgery is essential. In our institution this has been recognised and it has been taken forward as far as possible. However, a single institution cannot by itself transform a major specialty and the Department of Health, the Royal College of Surgeons, individual specialties, Regional Offices of the NHS Executive and other institutions all have responsibilities for achieving this. A concerted effort is needed to foster an effective R&D culture, including the very important debate about the stature, the credibility and the rewards of surgery in relation to research.

The real way forward is to target the young and to collaborate. We create multidisciplinary functional groups and theme interactions. We try to marry clinical excellence with academic strength, forming alliances outside GOS/ICH where helpful. Exposure to research principles; the pursuit of proper research; the protection of lecturers and career opportunities for our trainees; are all vital ingredients. The organisation and delivery of R&D is an academic discipline in its own right. It requires rigorous thinking, recognition of the fundamental importance of strategy, and tactical implementation. Applying these principles allows the welding together of teams and the reality of cultural transformation.

Discussion

PROFESSOR JOHN PRIMROSE

In discussing Al Aynsley-Green's paper, I would like to focus on two particular issues. Firstly, to look specifically at the NHS Trust environment from the perspective of a large, general, provincial Trust, rather than the London environment. Secondly, to look at R&D specifically from a surgical point of view.

The competing demands on surgeons' time have already been noted. Surgeons are very busy people and within a Trust they may have enough problems just getting routine surgical admissions processed through the hospital even before considering undertaking research. There is no question that there are cultural problems, especially in those Trusts which do not have a long tradition of being a teaching hospital. One must also recognise that, as far as surgeons are concerned, the culture of research as academics see it may not always be foremost and other career choices will after seem more rewarding. A further point concerns culture. One of our experiences in Southampton is that a RCT of a service provision, that was set up with R&D support, was regarded as a tremendous threat by surgical colleagues, because of their anticipation that it might take away patients currently managed by them.

The key issue is who pays in the current Trust environment. A paper by Solomon *et al.* (1994) published in Surgery a few years ago, looked at RCTs in surgery and found that a large proportion of them did not have external funding. This contrasts with trials of therapies, which are usually extremely well funded, usually by the pharmaceutical industry. There is no problem for a Trust in organising, or taking part in, trials of therapy, because the funding normally comes with the project. There is, however, no priority within a Trust to support surgical R&D.

My own experience is that regional R&D and other funding sources, such as the Health Technology Assessment Programme, will in fact support good quality, collaborative surgical research. They may, however, have to become more aware of the difficulties in designing sur-

gical trials, and perhaps be more pro-active in directing support to trials with pragmatic designs. The outcome of trials in surgical R&D may actually be of greater importance to the NHS than many currently performed trials of therapy, at least in financial terms.

With respect to the busy general NHS Trust, what happens when someone performs R&D and this leads to an increase in activity? No-one seems to be responsible for funding the increased clinical workload that such trials bring in. Certainly the Trust does not feel it is within their remit; the funding bodies are not prepared to pay; and the Regional Directors of R&D would say that it certainly is not their problem since the Trust has already got its 'Culyer' money designated to support such activity.

So, how can progress be made in the NHS Trust environment? Academic leadership needs to be strengthened and valued more than is currently the case within many Trusts. Perhaps Great Ormond Street is a model for this. The Culyer funding needs to directed to support the infrastructure for research active areas, rather than just being smeared across the Trust as a whole, like butter on toast. Surgeons need to promote collaboration with other disciplines more, because multidisciplinary research proposals stand a much higher chance of being funded. Lastly, we must all recognise simply that surgical research, particularly RCTs, can be very difficult to do.

REFERENCE

Solomon MJ, Laxamana A, Devore L, McLeod RS (1994). Randomized controlled trials in surgery. *Surgery*, 115(6): 707-712

Chapter 6
The role of purchasers in the introduction of new surgical technologies

DR. PAUL WATSON

My aim is to discuss the purchaser or health authority perspective on surgical R&D. The issues I raise apply to any innovation in health care, but the examples I use are from the surgical field.

Purpose and funding of R&D

Health authorities inevitably have a local focus. They are explicitly expected to look after the interests of their local population. By contrast, R&D in the UK is about advancing health technology for the benefit of the whole country, and therein lies the first dilemma. This leads to a situation where health authorities that host teaching and research centres have to make a trade-off between investment in developing new services and keeping the services that are currently funded on track.

This kind of trade-off has been greatly helped by the Culyer reforms to the financing of NHS R&D, where there is an explicit, separate funding route for R&D and the trade-off between investment in innovation and investment in service is made, quite appropriately, primarily at the national level. This does not, however, address the whole problem. Only a proportion of R&D costs is funded through the Culyer mechanism. Some R&D costs have to be met locally and this is a particular issue in surgical R&D as national funding streams for R&D are particularly poor in surgery. There is no pharmaceutical industry funding for such R&D and the MRC estimates that only around 15% of its national R&D funds are spent on surgery.

Surgical services: purchasers' concerns

When Trusts and individual clinicians ask a health authority to consider funding a new innovative service within surgery there are several factors that will influence the purchaser's decision. Cost control, not letting costs run ahead of available funding, often comes first because it dominates decision-making within a health authority, where there is a constant threat of the money running out.

Cambridge and Huntingdon Health Authority had a relatively small deficit in 1996/97, about 1% of turnover, but that had to be recovered in 1997/98. As a result, health visiting and chiropody services were cut and a range of elective surgical procedures were stopped. Anyone having had to go through this process would not wish to do so again.

Cost control is not the only concern, though. Service quality is important as well. The way that a health authority thinks about service quality is slightly different from the way that individual clinical teams often think about it, however. A health authority is formally an agent of the Secretary of State for Health. It is therefore beholden to implement the policies of the elected government of the day. One of the current policies, and therefore one of the concerns of health authorities, is to reduce waiting times or at least control them. Furthermore, access to emergency services is a particularly topical issue and tends to be particularly difficult in intensive care. Access to an emergency admission when it is required is one of the things that we are expected to guarantee to our local community.

There are a number of things that are of concern to health authorities when looking at the clinical quality of a contracted surgical service. These include staffing levels, team working, audit arrangements, skill levels and, lastly, access to new technology. The reason I put new technology last is that there are a number of areas where simply better coordination and organisation of the current NHS could yield better clinical outcomes. The Calman-Hine policy on cancer services (Department of Health and Welsh Office, 1995) is predicated on the assumption that a lot more can be done with current cancer treatments. For example, if there is greater sub-specialisation among surgeons, or better team working between surgical teams and oncology teams, this might improve patient outcome.

The fact that health authorities are dominated by cost control and have a slightly different perspective on clinical quality than clinical teams, means that there is often a clash of cultures when health authorities and surgical teams meet to discuss service developments. The possession of the latest diagnostic kit, or the ability to carry out

the latest surgical procedure, is not viewed by health authorities as a measure of success in their local surgical services. Quite understandably, individual clinicians want to be on the leading edge of development in their particular field, and will therefore take their ability to provide a new service as a measure of their own success.

A further reason why new technology creates a concern for the health authority is that it is one of the drivers of increased cost in the NHS. New technology can drive activity and therefore costs in the NHS in different ways. Clearly, the introduction of a new technology is an important decision to make in an individual Trust or health authority. The decision to introduce endometrial ablation for the first time in a hospital, or the decision to repair hernias laparoscopically for the first time, are important decisions which need to be properly managed. Although rising activity and costs in any local health care system are partly due to the introduction of technologies, the wider application of techniques which are already well established is also a big factor. As an example, in Cambridge and Huntingdon the operative intervention rate for lung cancer roughly doubled between 1995 and 1997. This was not as a result of the introduction of new operations, nor of deciding that there are new indications for operating in lung cancer cases. It was simply because the assessment process had been streamlined, so the referral threshold went down, patients got to the right clinician quicker, and activity went up. Nobody had made the decision that lung cancer was to be treated in a different way. What was happening was simply the more systematic application of an existing technology.

An interesting case study is percutaneous transluminal coronary angioplasty (PTCA), where there are two different types of cost driver. The first cost driver is the introduction of a new technology, namely coronary stents. This is a fairly new procedure that was not available a few years ago and is more expensive than previous practice. But at the same time the gross rates of coronary angioplasty are going up throughout the NHS. This is due in turn to changes in referral habits and a widening of indications for PTCA. Cardiologists are more liable to intervene in a wider range of problems and over a wider age group than they were ten years ago.

The process of introducing new surgical technology

The introduction of a new procedure in surgical services will inevitably be vying for attention within a health authority with a lot of other issues that will be bubbling up at any time. But it is nevertheless a process that needs to be managed. The process that a health authority would go through when thinking about introducing a new service, whether a surgical procedure or a drug or something completely different, requires the demonstration of four aspects (Figure 6.1).

Figure 6.1 **The process of introducing new surgical technology**

The first issue is about the efficacy of the procedure, and this is simply to consider whether the technology can work, even in principle. In the very rarefied atmosphere of an RCT, has the technology been shown to benefit patients? The relevant evidence is not always an RCT, although clearly that is the gold standard and should be used when feasible. There may be other types of evidence to use as well, particularly observational trials.

If the technology is shown to be efficacious, then an economic evaluation needs to be performed, comparing the incremental cost of introducing the new service with the incremental benefit that it will produce.

The next issue is about effectiveness. Efficacy and effectiveness are often taken to be the same thing, but there is a clear distinction between them. Efficacy is about whether a technology can work in the context of a trial and it is a necessary but not sufficient condition

for effectiveness. Effectiveness is about whether the technology will work once it is introduced into the everyday NHS setting. There are a number of factors that will determine whether a technology that works in a trial will actually work once in practice. There is a generalisability problem since most RCTs have strict patient selection criteria. But once the new service is introduced to the NHS patients are usually selected from a population with a much wider range of clinical symptoms and other characteristics (such as age). Something that is proven to work for a particular indication in a particular age group may not work once it is used for other indications or other age groups.

There is also an issue of operator characteristics. In many of the trials the people who are allowed to perform the technology are strictly monitored. A good example are the carotid endarterectomy trials, where surgeons were not allowed to take part unless they were doing a certain volume of procedures per year and could demonstrate they were doing better than a specified acceptable level of complications. Once endarterectomy has been introduced into the NHS, this tends to be forgotten about and the complication rate or volume by operator is not systematically monitored. Without monitoring of these sorts of factors, we cannot know whether what was shown to be efficacious in a RCT actually works once it is introduced into the NHS.

Support services also influence effectiveness. There is a whole raft of support services which are important to providing a good clinical outcome for a particular service. It is not solely the surgical operation itself that is going to deliver the beneficial outcome. There are other factors to be addressed. There is the phenomenon that patients entered into trials in cancer tend to do better, whether they are in the control arm or not, than patients who are treated outside trials. It may be that the support process and the amount of attention that is given to patients in trials have an independent beneficial effect on outcome.

Lastly, and crucially for health authorities, the affordability of a new procedure is a key factor determining its introduction. A procedure could be efficacious, it could have a reasonable cost per QALY (quality-adjusted life year), it could be shown to be effective once it is in NHS practice, but it may simply not be affordable. The growth money

that the NHS receives each year has usually been insufficient to keep up with the legitimate demands produced by the health care system. Something has to give. What should it be? This is a dilemma that health authorities wrestle with all the time.

The scientific base for demonstrating the four aspects of the purchaser's decision making process shown in Figure 6.1 – efficacy, economic evaluation, effectiveness and affordability – diminishes the further down the diagram one goes. The first RCT was performed about 50 years ago and the general methods for RCTs are fairly well established, although there are still quite a lot of practical problems in performing RCTs in surgery. Economic evaluation has been around a while now. There have been some very robust studies carried out and there is a relatively good academic base there. With respect to assessing effectiveness we are still just getting past first base. We do not have mechanisms to systematically measure outcomes and appraise the translation of innovation into practice. It is not something that the NHS addresses at the moment to any great extent. The effectiveness/efficacy gap could be addressed through more systematic audit. The audit process at the moment is largely unmonitored by health authorities, and there is a clear need to improve this. In terms of affordability, we are not yet at first base. The implicit government policy is that the NHS is completely affordable: that it can absorb any demands thrown at it. So even saying that something is unaffordable is still regarded as heretical within the NHS.

Funding new surgical services

If the health authority decides that a new service is worth introducing, then there are basically three sources of money available to fund this. The first is through the growth money allocated by the Department of Health each year. This is usually around 2% extra each year, after inflation.

It is increasingly obvious that not all legitimate demands can be funded through new money, so the second option is to fund new developments by redirecting funds from other areas within the hospital and community health service budget. This option was adopted by Cambridge and Huntingdon Health Authority, where health visiting

and school nursing were cut and a range of surgical services stopped, in order to fund other services, such as emergency admissions, cancer care, and so on. One method of introducing a new service is therefore to demonstrate that it is a more pressing call on the local NHS budget than are other services that are currently already funded by the health authority.

The third option, and perhaps the most painful for surgeons themselves, is to reduce other surgical services in order to fund the new surgical innovation. If there is a programme budget where the health authority makes it clear that a certain amount will be spent on surgical services, then if a surgical innovation comes along that will cost more money, which existing services within the surgery budget should be reduced? Is it about being even more aggressive about day case rates? Is it about stopping doing certain types of surgery altogether? Is it about introducing selection criteria for various operations? These are real questions that health authorities are increasingly having to answer.

Refusing new surgical services

Alternatively, a health authority may decide not to fund the new service. There are a several grounds on which this decision is made. The first is that a service is unproven. There are difficulties with this line of argument since a lot of the services that are already being funded at the moment are unproven, and probably never will be proven.

The second is on the basis that the service is ineffective; that is, there is good evidence that it simply does not work. There is a big difference between the unproven category and the ineffective category. The unproven category is really saying that there is an absence of evidence of clinical effect, whereas the ineffective category is saying there is evidence of absence of clinical effect. An example in the unproven category would be lung reduction surgery in emphysema: nobody knows whether it is of benefit or not. Clearly, the appropriate way forward is to address this through a trial. In the ineffective category, if somebody came up with, say, laser tonsillectomy for glue ear, clearly that would be nonsensical because we know that tonsillectomy cannot work for glue ear.

The third ground for turning down a new service is that it is an inappropriate use of NHS funds. There are some areas where there is general agreement that it is inappropriate to fund. For instance, spending NHS funds on providing face lifts. But there is often a grey area where some argue that something is an appropriate use of NHS funds and others disagree. In the end this is a value judgement.

The final reason that can be given for refusing to introduce a new technology is that it is unaffordable. It may be worthy of NHS funding but it simply cannot be afforded.

One of the problems that health authorities often get themselves into is that they muddle up these four categories. The first two, the unproven and ineffective categories, are primarily scientific judgements about whether a service has a proven basis or not. The other two are value judgements. They are political, moral and ethical decisions about how the community should fund its health services. Health authorities are often very uncomfortable with making these kinds of moral and ethical decisions, and it is all too tempting to hide behind science as a way of making what are in the end ethical decisions.

Obviously, honesty on the part of the health authority is required in arguing against a technology using these categories. It is dishonest to say that a given service is not to be supported on the basis of spurious science, which is why we need to be very careful about the misuse of health technology assessment by health authorities. It is perfectly acceptable to decide that a service will not be introduced because it is ineffective or unproven but it is not acceptable to pretend that the grounds for rejection are some spurious health technology assessment when the real reason is unaffordability. It is also not fair on the people who do the assessments in the first place to misuse their work in this way.

A managed process for the introduction of surgical technologies?

At present we have an unmanaged process for the introduction of new technologies. A 'licensing process' for the managed introduction of new surgical techniques into the NHS is required. There are a num-

ber of areas of very good work, such as the work of the HTA programme, NICE, SERNIP and so on. Yet there is no systematic way to turn that information into decision-making on whether and how services get introduced. There needs to be and it needs to be transparent. It is not fair on surgeons, it is not fair on patients, and it is not fair on payers, whoever they may be, to have the current opaque system.

The very good work that has been produced through the HTA programme has to stand as objective, scientific appraisal. It must not be misused in order to dodge the rationing issue. Systematic work on the gap between efficacy and effectiveness is required. Lastly, honesty on affordability is needed. Although investment is made in R&D to push forward surgery in the UK, this does not necessarily mean that the outcome of that R&D can necessarily be afforded by the NHS. Unless we have honesty, we are being unfair to everybody.

REFERENCE

Department of Health and Welsh Office (1995). *A Policy Framework for Commissioning Cancer Services: a Report by the Expert Advisory Group on Cancer to the Chief Medical Officers of England and Wales.* London, Department of Health.

Discussion

PROFESSOR RICHARD LILFORD

I want to draw a distinction between two kinds of research funding: explicit and implicit. We are most familiar with hypothecated research studies where the funder provides resources explicitly for a particular study, which runs for three years or so, ends at a particular date and produces results further downstream. I am not against that mode of research funding but I propose that there should also be capacity provided for trials to develop in an organic way, where researchers do not have to wait for the grant giving bodies to decide whether to fund. Such implicit research allows studies to start rapidly when clinical questions arise.

In terms of hypothecated forms of research, there are three kinds of costs that need to be financed. Firstly there are research costs, which are funded from the Health Service Research Board of the MRC or some other grant giving body. Secondly there are the support costs, which cover things like the extra clinic time taken to explain to a patient the test or treatment options available and hence whether to enter a clinical trial; or extra tests, such as MRI scans, required purely for research purposes. Trusts bid for these costs from the NHS R&D programme Budget 1. The third type of cost is the treatment costs, which are the net excess costs of the actual treatment invoked in the trial. Such costs would continue after the trial were the new treatment to come into routine use rather than the current treatment.

Purchasers (health authorities) have no option but to meet these treatment costs. The government has said, via its guidance note HSG(97)32, that the existing funding mechanisms in the NHS must meet these costs locally. If this had not been done then purchasers might have spotted an opportunity for more money by labelling services as research. For that very good reason the government decided not to take these costs out of the normal local patient care funding process.

The issue of how purchasers should deal with the treatment costs remains less clear cut for non-MRC, non-NHS R&D funded pro-

grammes, however. In these circumstances, a health authority can seek to have the project re-assessed in terms of value for money if excess treatment costs arise. There is also a second escape valve for receiving additional funding and this applies to all studies, whether funded by charity, by the MRC or by anybody else. If a study is particularly expensive compared to the generality of trials carried out at your local hospitals on a regular basis, then a purchaser can ask the NHS Executive for a subvention – extra money to meet some of the extra costs which arise as a result of hosting the study. An example of this is Dr. Alan Scott's trial of screening for abdominal aortic aneurysms. Firstly, the screening itself costs a lot of money. Then, if an aneurysm is found, the patient goes to theatre and has it operated on. This generates high costs and no hospital could bear them without disrupting normal patient care. Another trial for which a subvention has been available concerns cerebral artery aneurysm treatment. Such subventions apply only to studies in which there is an external funder. They do not apply to a study funded within a hospital.

Regarding external funding, I am unhappy with the general concept that information about the health service resource use of the new technology should generally be collected within effectiveness trials. What I think should instead happen is that the economic evaluation should be performed before the clinical trial is started. In other words, a pre-trial model should be built. The purpose of the trial would then be to create more precise data to help populate that model. One advantage of this approach is that it may establish whether a trial is actually needed in the first place (i.e. you already have sufficient information without a trial), or whether the issues to which the subject is most sensitive do not lie within the trial domain but can be collected elsewhere. For example, a trial has been mooted of how to manage mildly abnormal cervical smears, where the options are to repeat the smear six months hence, or to refer for immediate colposcopy. A pre-trial modelling approach was used and found that the costs turned on the threshold the laboratory had for recording the test as abnormal because this determines the chances of the subsequent smear still being abnormal and the patient requiring colposcopy (Johnson *et al.*, 1993). So to do a trial at one level of

reporting would only answer the question at that level of reporting, not for all the other possible levels; and we know that around the country laboratories vary in how they report.

There is another issue which is often not addressed when considering the economic consequences of a technology, and that is in whose gift the technology finally lies. There is a spectrum with at one extreme the technology being in the gift of the funder of the service. For example, there is a trial of Extra Corporeal Membrane Oxygeneration (ECMO), an artificial lung for very small babies, which is a very expensive technology. Clearly, a clinician cannot just go and get some ECMO. First, they have to get hold of the ECMO machine and a place to put it, and then find the nurses and other staff to operate the service. That cannot happen unless the managers have agreed to proceed with the technology. At the other end of the spectrum, there are treatments entirely in the gift of the clinician. I am doing a trial with colleagues (Thornton and Levine) on what to do about the very small baby who is failing to thrive inside the uterus. Should the baby be delivered early and put in an incubator, or should it remain in the uterus for as long as possible? This is a choice which can only lie in the individual consultation room. An obstetrician can, after consulting with the patient, decide to deliver the baby, creating huge consumption of health service resources in consequence. In such a scenario, the question needs to be considered, before spending a lot of money collecting health economics data, whether it is going to make any difference to the clinician's decision.

Modelling in advance has the further advantage of determining scope. Epidemiology needs to be considered before planning a trial, especially a hypothecated trial. It will tell you what the technology *could* do. If funding for trials has to be prioritised, then the potential gains from each technology have to be considered. This is often a rather depressing statistic from the point of view of my previous specialty of obstetrics where, in terms of screening for obstetrical conditions, the potential gains (in terms of epidemiological plausibility) are small. As a result, a trial of, for example, abdominal aortic aneurysm screening is more likely to win funding.

A further advantage of pre-trial modelling is that it can address the issue of generalisability. In terms of entry criteria, the usual practice is to have inclusive, not selective, entry criteria. A good example arising from surgery was the carotid endarterectomy trial where there was much disagreement about who should go into the trial. The trial was therefore based on equipoise: in other words, when the patient and doctor together were resting on a fulcrum of a decision, that was the entry criterion for the trial. Stratification then produced quite a clear answer: severe stenosis is best treated surgically, moderate stenosis medically.

A further point about the design of the economic evaluation of a technology is that a lot of the costs depend on results quite remote from those actually measured in the trial. The classic example of this is the study of a very tight versus 'ordinary' glucose control for insulin-requiring diabetics, where less renal disease and eye disease were found when tight control was used. The major economic consequences start from these outcomes. The savings from not having to treat the social consequences of blindness and not having to provide treatment for end-stage renal disease are great in comparison with the costs arising in the trial itself.

REFERENCE

Johnson N, Sutton J, Thornton JG, Lilford RJ, Johnson VA, Peel KR (1993). Decision analysis for best management of mildly dyskaryotic smear. *Lancet* 342: 91-96.

Chapter 7
The role of the Royal College of Surgeons

PROFESSOR PETER BELL

Introduction

Surgical research is little different from any other kind of medical research except for important differences in surgical technique and expertise, which can vary between investigators. Surgical research is driven by individuals, and surgeons tend to try and do things themselves, driven by industry and changes in techniques. Unfortunately, trials are rarely done. The reasons for that are numerous and have been discussed in earlier papers. So how do we actually try and advance research in surgery without suffocating people with too much bureaucracy? It is hard to answer this problem, but I will describe how the Royal College of Surgeons is attempting to promote surgical research.

Are surgeons guilty of not assessing new techniques properly?

Horton produced an editorial in The Lancet (Horton, 1996) that was highly critical of surgical research. This was prompted by the publication of a particular study (Majeed *et al.*, 1996), which showed that, in a properly randomised investigation, results from laparoscopic cholecystectomy were no different from those obtained when a small incision was used, except that laparoscopic surgery was more expensive. This was the first time a trial had been done to compare laparoscopic and open surgery. The results have, however, been completely ignored and surgeons go on doing laparoscopic surgery in spite of its greater cost.

The rules for performing trials are well known. Sackett (1989) identified these a long time ago and categorised types of trials in descending order of reliability from Level 1 to Level 5. Many of the studies that surgeons are involved in are at Level 5, the lowest, i.e. 'I did a couple of cases, they did very well thank you very much and now everybody else should do them'.

A Consort statement was published (Begg *et al.*, 1996) which discussed randomised trials and proper reporting standards. Randomised trials are interesting in that if one looks at them carefully they are, in practice, often not randomised correctly. The Consort

document states that randomised trials should report which patients were left out and why. Often, patients are excluded for reasons purely of convenience. So, are surgeons guilty of not assessing new techniques properly? Yes, I think they are.

RCTs are not the only useful form of assessment, however. Descriptive papers are important too. It is essential that a new technique which has evolved, possibly from just one surgeon, is described and it is very important that, after this, a feasibility study is done to prove the worth of the technique. Various registries are now emerging which list the number of cases done when these new techniques are used, and comment on how badly or well they are done and what the complications are. After the feasibility study has been completed, it is then reasonable to go on to a randomised trial.

Encouraging surgical research

Figure 7.1 lists the key requirements for encouraging surgical research.

The Royal College of Surgeons and others have to promote and stimulate research and its assessment. Unless an interest and awareness in research is fostered, surgeons will in general not do any research. The College has said that research is desirable in the education of all surgeons for a period of about one year, full time. The College backs up its words by providing research fellowships, by supporting professorships and lectureships, and by encouraging surgeons to travel between centres – nationally and internationally – to look at research. It also supports scientific meetings. Overall, the Royal College of Surgeons makes a considerable research commitment.

Figure 7.1 **Key requirements for surgical research**

What we need is –

- Venture capital
- Support of grant giving bodies to recognise the importance of surgical research
- Time and salaries for young surgeons to do the research

When is the best time to do research in surgical training?

In surgical training, the main question is when research should be done. There is a tendency for surgical trainees to do their research between the MRCS (basic surgical training) and higher surgical training, the latter lasting for five or six years. It is debatable whether or not this is the best time to do research. Many surgeons have, in the past, been turned off from doing research because they feel that they have been made to do it in order to get a job later on. That should not be the case. Surgical trainees should do research beyond a limited period only if they want to. Those wishing to do research for longer periods, particularly those wishing to embark on an academic career, have problems however. If a thesis is the aim, then that means that the aspiring academic has to do two or three years extra work above and beyond the training done by others. At the other end of the spectrum, merely providing surgical trainees with casual involvement with ongoing research projects is not ideal, but it may at least foster the research habit in the trainee.

More recently the MSc has been introduced in many institutions. The Royal College of Surgeons is itself about to launch an MSc. This is a shorter degree allowing trainees to be introduced to research methodology with a short project. Whether this type of research will catch on remains to be seen. Most interested parties take the view that part time research is not really much use. Those who want to do serious research need to spend the whole time doing it.

Funding surgical training

The next problem of course is how to fund the research. There is basically no funding for a research year during the Calman training programme, so alternative sources have to be found. Money can be obtained from various organisations such as the British Heart Foundation, the MRC and the Wellcome Foundation, but these institutions often want a named person before considering giving a grant. Such a person has to be employed somewhere, a process which can take a year and it is very difficult to have an applicant waiting for that length of time. The answer is often to use soft money or involve private hospitals in financing research fellowships. The College, fortu-

nately, has now stepped in to provide money to bridge this gap. The College has used its funds to provide a number of ways of helping with research and has produced a series of brochures to advertise the fact. Professor Aynsley-Green mentioned earlier the vital importance of marketing and presentation, and he is right of course.

The one-year research fellowships offered by the College are meant to be an opportunity for young surgeons doing the Calman training programme to do research to a reasonable standard. The applicants apply in open competition and their applications are peer reviewed. The project, the applicant and the supervisor are all judged. The supervisor is particularly important because of the criticism one often gets from non-surgeons that research and higher degrees in surgery are of a low standard because of a lack of supervision. Unfortunately this is often true. My own first experience in research was when I was appointed as a senior house officer in Sheffield, given a desk and told to get on with the research. I did not know what to do, so I thought I would find a cure for cancer! Of course, that is what one tends to do when one first starts doing research. Supervision and guidance are therefore absolutely vital.

There are also three-year training fellowships in association with the MRC, although only a few each year are granted at the present time. These are important for people wanting to do serious research in order to gain a PhD or an MD. Applications for these are assessed in the same way as for the one-year fellowships, by peer review.

Apart from fellowships, the College provides pump priming grants for newly appointed consultants and senior lecturers. Someone who is keen to do some research gets appointed to a consultant post and finds that there is no money to start research. He then goes into private practice and loses the urge to do research. The College therefore provides pump priming grants as an opportunity for the candidate to obtain up to £20,000 to start their research.

The joint Royal College of Surgeons and MRC training research fellowships mentioned earlier, are an important advance. In the past, the MRC had been seen as not funding surgical research and these joint fellowships are a particularly good way of rectifying this. It has led to

a lot more surgeons applying for research grants with the MRC, who are now more aware of the quality of surgical research.

Finally, the College provides further joint fellowships in combination with local charities which share the financing.

Between 1993 and 1997, the College handed out over 100 fellowships and spent £2.5 million in doing so. This money was previously used to run research departments which were working in isolation, and were not doing as well as they could do because of that. I think the College has been very successful in encouraging research. As a result of this success, other Colleges have started to do the same thing. For example, the Royal College of Surgeons in Edinburgh now offers similar fellowships.

The College makes efforts to publicise what it is doing in research. The successful candidates for College research fellowships are encouraged to attend open days and talk to people about their research. They also have their photographs displayed in the College portals, which is all good marketing. Graduates see this when they come into the College and they think 'I'll try and get one of those.'

A booklet is produced and sent out to all the Fellows of the College, telling them what has happened. Inside that is a report from each research fellow himself, and on the back is a list of all Fellows who have donated money to the College. This is not meant to be in any way threatening to those who have not made a donation, but it does seem to have a certain effect. In other words, the College's Fellows themselves donate money to this programme, and in 1996 about £70,000 was raised in this way.

Audit

The College, apart from getting money for research, is also involved greatly in audit and epidemiology and this has been going on for some time. Brendan Devlin started off the whole process and in 1998 the College appointed a Director of Epidemiology and Audit in collaboration with the London School of Hygiene and Tropical Medicine. The College has also appointed an assistant director, a surgeon working part-time, who is responsible for liaising with other surgeons and

raising the questions that need to be answered whilst encouraging their support.

Ensuring proper assessment of surgical techniques

Money is not always available to look at new techniques but the Health Technology Assessment (HTA) programme has helped in this direction. I am afraid, however, that it does not always help surgery greatly. Important questions such as 'does this operation do better than that operation', seem to come a long way down the list of priorities. In Holland there is a system financed by the government where applicants can pose a question and apply for support after appropriate peer review. Such a system should exist here too if we are to encourage surgical research in the future.

An important question is how to ensure that a technique is properly assessed before it creeps into general and usually unproven use. This is a very difficult problem but there are various ways of tackling it (Figure 7.2). Purchasers should not buy the technique unless it is part of a proper evaluation study. Alternatively, the medical defence unions might refuse to support surgeons who use techniques which have not been proved. Private medicine is often the engine room of unproven techniques and so it may be that the defence unions should consider what to do about this, along with the health care providers.

There are plenty of examples of surgical techniques that have been introduced without proper evaluation beforehand. Aortic aneurysm stenting, which is currently popular in vascular surgery, is an exam-

Figure 7.2 **Ensuring proper assessment of a technique before it is used generally**

- Purchasers shouldn't buy it, or only where it is part of a proper study
- The Defence Unions refuse to support the surgeons unless the technique is part of a trial
- Private medicine will not buy it until its efficacy is proven
- Money must be provided to assess it

ple of how things should not proceed. The usual treatment for aortic aneurysm was to open the patient and insert a graft. A doctor called Palmaz invented a metal stent which can be expanded with a balloon to stay in place inside an artery. A graft can be stitched onto the stent and placed inside the aorta to be used as a variant of keyhole surgery. This technique is attractive to the public because it is less invasive, but it cannot be used on everybody because the anatomy required to retain the stent has to be precise. The whole device is expensive. Various studies have now been reported, showing no increased mortality associated with use of the device but demonstrating problems with it, such as leakage. In this position one would think that trials should be done, but they have not been and industry, which is in for a profit, is happy to drive the process forward. Each meeting one goes to contains a number of different companies all trying to sell new stent/graft combinations for the treatment of aortic aneurysms, but there are no data to support their use. I am glad to say that a properly funded, randomised study has now started.

The results of any new technique can be made to look good if there are no controls. So why are surgical trials not done? A view is often heard that technology changes so rapidly there is no point in evaluating the current equipment because it will improve. A greater problem is that not enough people can be trained to do trials properly. Furthermore, if a trial is done, someone has to pay for the equipment as the manufacturers will not provide it free. This does not happen in drug trials but pills are a lot cheaper than stents, which can cost between £3,000 and £4,000 each. Who pays for this? Not the MRC, they would not use public money to pay a company for profit. The answer is not easy.

Promoting the diffusion of new ideas in surgery

The College is trying to help diffusion by providing hands-on courses where experts are available to instruct the would-be user on how to use new techniques and equipment. This will become more important as post-qualification medical education becomes mandatory. NHS Trusts should encourage people who are doing new things to attend such courses and should provide them with funds to do so. For

Figure 7.3 **The role of the College in promoting the diffusion of new ideas**

- In-house journal and bulletin
- Supporting meetings
- Via the examination process
- Providing courses

example, surgeons can thereby be shown laparoscopic surgical techniques properly, rather than testing their skills, or lack of them, on a patient before they have practised on a model. The College also has a journal and supports meetings to inform surgeons about new ideas. It has an examination process and provides courses. Through all these functions it tries to support surgical research and make it part of training. The role of the College in promoting the diffusion of new ideas is summarised in Figure 7.3.

The future

One way to proceed in the future is first to establish on a more formal basis that surgeons do get some research training, if only for one year, and money should be provided to make this happen. It should also be possible and a matter of priority to provide funds to do randomised multi-centre studies in any new technique that is used in the future. The way in which laparoscopic surgery and endovascular procedures are being introduced is an example of how things should not be done. Unless money is made available for studies, surgeons will never change. Simply saying that surgeons should not do new procedures is not a way forward as this stamps on innovation and prevents change, and I am sure that no-one would want to see that happen.

REFERENCES

Begg C, Cho M, Eastwood S, Horton R, Moher D, Olkin I, *et al.* (1996). Improving the quality reporting of randomised controlled trials: the CONSORT statement. *JAMA*; 276: 637-639.

Horton R, (1996) Surgical research or comic opera? Questions but few answers. *Lancet*; 347: 984-985.

Majeed AW, Troy G, Nicholl JP, Smythe A, Reed MWR, Stoddard CJ, Peacock J, Johnson AG (1996). Randomised prospective single blind comparison of laparoscopic versus small incision cholecystectomy. *Lancet*; 347: 989-994.

Sackett DL (1989). Rules of evidence and clinical recommendations on the use of antithrombotic agents. *Chest*; Feb 1989; 95(2 Suppl): 25-45.

Discussion

PROFESSOR ARA DARZI

My first point is that surgical research may be divided into so-called mechanical technology and biological technology. Unlike academic departments of surgery in this country, if you look across the Atlantic, over the last decade or so most of the funding has tipped the balance towards biological technology and biological research. I find that sad. Rather we should be trying to get people to fund research into mechanical technology. This is really where the role of the Royal College is important in funding research, rather than the traditional funding that is received from the MRC or the Wellcome Trust that is usually in the basic science, biological area.

As far as training is concerned, I suspect most surgeons are always trying to excel in their surgical rather than their research prowess. This can lead to problems when capable but non-academic surgeons try to evaluate a new technology. So the Royal College's role in setting up a training programme is important. The training facilities in the College are supported by the Wolfson Foundation and the Department of Health. The state of the art surgical training skills laboratories provided will arm any surgeon before entering the operating theatre with the basic skills required.

One of the questions that arises from training is who should fund the training programmes in the newer technologies? That has been a major issue. One does not expect a trainee to pay an average £2,000 a year to learn three or four procedure-specific courses for new technologies, such as the insertion of a vascular stent in an aortic aneurysm, for example. These are expensive courses, and the resources need to be found to fund individuals to attend them. Otherwise, high quality courses will exist but there will be no mechanism of ensuring that surgeons are fulfilling the basic requirements in training before they undertake a new procedure. I do not really know who needs to address that: whether it is the purchasers; whether the Trusts should pay for their surgeons to attend courses at the Royal College, or any of the courses that are franchised to their own Trust, before they embark on a new procedure; or whether it should be the government who will provide a substantial amount of money for training.

Chapter 8
Summing up

PROFESSOR SIR MILES IRVING

Is surgical research different?

Fundamentally, surgical research is no different from any other research that involves the treatment of patients. I think all of us would agree, however, that it is more problematic. Having been involved myself in a number of RCTs, organised by others on a multi-centre basis, I know that one of the issues is preventing the trial workload interrupting daily work. Can anything be done to improve the situation? I think the answer has to be 'yes'; and I think the first thing to do is stop ourselves thinking that service provision prevents us from conducting research.

Surgeons' interest in evaluation

The old story of 'problem surgeons' is a bit of a myth, in my view. Surgeons have not always done as well as they could with evaluation, but at the same time there is not a lot to be ashamed about either. More than 200 years ago, John Hunter (1728-1793) pointed out that if surgeons were going to be good surgical scientists they had to follow up, closely record and learn from, the progress of patients. In retrospect, the individual who had the most impact in this area was Ernest Codman (1869-1940; see for example: McLendon, 1990, for a review of Codman's contribution). He made it quite clear that the monitoring of outcomes and the methods of introducing new procedures should be systematised and orderly. SERNIP's regulations and description of how a new procedure should be introduced are virtually identical with what Codman recommended. Interestingly, Codman then said that, having done the efficacy studies, the next move was up to the College, in his case the American College of Surgeons. He considered that it was the College's job to supervise the dissemination of the new technology. I will come back to this because I think it is still relevant.

There have been some successful randomised trials in surgery in the UK. Goligher performed one of the earliest randomised trials which

gave very clear results affecting treatment (Goligher *et al.*, 1968). The MRC European Carotid Surgery Trial is a recent good exemplar (European Carotid Surgery Trialists' Collaborative Group, 1998). In my own speciality, colo-rectal surgery, I find evidence-based guidelines for the management of colo-rectal cancer, produced by the Royal College in association with the Association of Coloproctology. There are several RCTs and a cost-effectiveness analysis with a positive outcome; and there is a Cochrane group associated with colo-rectal cancer. Given the opportunity, surgeons are indeed interested in evaluation. It is other factors that cause the problems.

The scope of surgical research

Everyone would agree that there has to be collaboration between the three major research areas: the laboratory; the technology transfer from laboratory to efficacy studies; and using HTA to address effectiveness in use. The best laboratory research in the world is useless if it does not then make the transfer to efficacy studies. It is therefore very important that surgeons working in collaboration make the transfer from the laboratory to the specialist centre, where the technology is investigated in clinical surroundings, and then onwards to practice.

The MRC funds the laboratory work and the surgeons themselves, through the Colleges and the charities, fund most of the research fellows at that level. The transfer from laboratory to efficacy study can now be funded through the NHS R&D mechanism, sometimes by regional R&D but also by national R&D. The HTA programme funds this aspect as well.

Evaluation is a key part of health services research and should not be bypassed. Virtually all effective evaluation was bypassed in the case of laparoscopic surgery, where a new technique was introduced driven by commercial pressure, personal enthusiasm and patient demand. So students have to be educated, both at undergraduate and postgraduate level, that evaluation is important. Purchasing authorities also have a role to play in encouraging evaluation. They have to say that if the NHS, or the government through the MRC, is investing in an evaluation, then it is quite wrong for other surgeons to weaken that eval-

uation by using the new technology without incorporating it into evaluation studies.

The NHS Executive's 'Executive Letters' to health authorities do provide support to the purchasers in this matter by saying that if a procedure is still being evaluated then it does not recommend that it be purchased. Sadly, sometimes such procedures are still purchased, although the Culyer funding mechanism may reduce the likelihood of this happening in the future. If some NHS Trusts are deliberately undermining studies that are being carried out and funded by other parts of the NHS, then the threat of discontinued Culyer funding to that particular Trust could be a very powerful incentive to comply. This mechanism has already been used in that way.

Problems of *evaluation*

There are several problems in designing evaluations. There is a culture that makes evaluation difficult. It is not just the profession, it is the public as well. I think a major effort has to be made to influence the public. About half of my teaching time is now involved in giving lectures at the request of the public. I spent a day talking to 700 members of the Multiple Sclerosis Society and not one hostile comment about our evaluative approach to the use of beta interferon emerged. Indeed there was deep understanding of the need for evaluation. Now, of course, that is a selected group but if they can understand then I think the public as a whole can understand. I think we must also get on board professional help in getting the public to understand two concepts: one is 'evaluation' and the other is 'uncertainty'.

There are 'problems with RCTs but Robin McLeod, Professor of Surgery in Toronto, has pointed out (McCleod *et al.*, 1996) that if you look at surgical techniques, about 40% are suitable for randomisation. Excuses can be made for not investigating conditions eminently suitable for test by randomisation but there is no justification for that. There remain, however, 60% of conditions where other types of study than RCTs will have to be performed.

The problem of the learning curve in new surgical treatments is an issue that the Methodology Panel of the HTA Programme is address-

ing. There are methodological scientists amongst us who say that the learning curve can be incorporated into the assessment of a technique. I do not play down the difficulties in such trials, but I think there are solutions.

RCTs or HTAs cannot be performed on absolutely everything and other techniques are required. Once again, the Methodology Panel of the HTA Programme is looking at this. I have no doubt in my own mind that the DEC methodology pioneered by the Wessex Institute for Public Health (see chapter 4 above) is a way forward which will give purchasers and providers a quick answer. It is encouraging that a lot of these quick answers are confirmed as correct by other studies later on. We have to develop a mechanism for increased collaboration with that approach.

Concerning money, I think that if we use the existing funding properly, then enough is already in place.

Opportunities

I think the most important opportunity is that the NHS has great potential to be used as a test bed. One of the problems that I see with surgical trials is that, with poor recruitment rates, they can take a long time. Surgeons have got to learn to use the NHS as a test bed and with 60 million people in this country all using the NHS, the opportunities to answer questions quickly if everybody collaborates is unique.

So once again we return to how to encourage the surgical profession to take part in research. The answer in my view must come through our professional bodies. I think the Colleges have made a good start and they ought to carry on and take the agenda further. Trainees need to be educated of the necessity for evaluation.

Within surgery there are specialty associations for particular aspects of surgery. The colo-rectal surgeons have got off to a good start in this respect, and have produced evidence-based guidelines. The urologists have followed and others need to be brought on board too. I think also that there has been a change of hearts and minds in many areas, and it is quite interesting at surgical meetings now to hear people stand up questioning the evidence base for a particular technique.

The other opportunity is that there is now an early warning system in place, identifying new technologies, which needs to be made very effective.

Solutions

The four 'P's are a start: we need to influence the profession, the public, the press and the politicians. Sadly, I think the fact that annual NHS R&D funding has recently been cut by £10 million, although not a big amount in relation to the total, is perhaps big enough to send a message that the politicians have not yet been fully convinced of the value of evaluation.

There have been some very successful trials that have been carried out and the reasons why these were successful need to be explored. We need to look at new methodologies: explanatory trials versus pragmatic trials. Modelling provides an interesting framework for evaluation but it needs an evidence base first.

Rewards for undertaking research are very important. One option that needs to be explored is giving time, rather than money, as a reward: time for surgeons to perform operations that are going to be incorporated into studies; and also time to think. If there is going to be continuing professional development, colleagues in the NHS need something like a sabbatical at regular intervals. I am not talking about a year's sabbatical but short periods to think, to learn, to train. I believe the Royal Colleges ought to be a focus for training, and also for the advancement of the health services research in general. The concept of the MSc in health services research currently being put forward by the Royal College of Surgeons is a very good way forward.

The involvement of industry in the so-called 'fourth hurdle' of establishing the cost-effectiveness of a surgical technology is important. The creation of NICE in April 1999 has provided a strong stimulus for that. Industry has been quite fascinated by the whole R&D process but also somewhat disturbed by it. Nevertheless it too now sees the necessity for being involved in an evaluative process.

Future directions of the R&D programme

HTA asks 'does it work, for whom, at what cost?' But there is another research area where surgeons can be particularly involved, and that is service delivery and organisation. So, having asked 'does it work?', the second question for research is 'how can it be delivered?' How surgical services are delivered is something which surgeons ought to be thinking about already. In fact, the Colleges have already produced documents about the delivery of services. What is needed now is an evidence base. Service delivery and organisation is a new research programme, with funding started in 1998.

A smaller but also important area is evaluating new interventional technologies, in other words those where industry does not have much interest in providing the technology concerned, but the NHS sees a need for it.

Overall, what is needed is to bring more people on board into the whole area of evaluation and to use the NHS as a test bed, for it offers a tremendous supportive structure to allow us to undertake these evaluations. I think if we do this, and if we are led by bodies like the Royal College of Surgeons, then holding this meeting in another ten years time might tell a very different story.

REFERENCES

European Carotid Surgery Trialists' Collaborative Group (1998). Randomised trial of endarterectomy for recently symptomatic carotid stenosis: final results of the MRC European Carotid Surgery Trial. *Lancet* May 9; 351(9113): 1379-1387.

Goligher JC, Pulvertaft CN, De Dombal FT, Conyers JH, Duthie HL, Feather DB, Latchmore AJ, Shoesmith JH, Smiddy FG, Willson-Pepper J (1968). Five to eight year results of Leeds-York controlled trial of elective surgery for duodenal ulcer. *British Medical Journal* June 29; 2(608): 781-787.

McCleod RS, Wright JG, Solomon MJ, Hu X, Walters BC, Lossing A (1996). Randomised controlled trials in surgery: issues and problems. *Surgery*, May 1996; 119(5): 483-486.

McLendon WW (1990). Ernest A. Codman, MD (1869-1940), the end result idea, and The Product of a Hospital. The challenge of a man ahead of his time and perhaps ours. *Arch Pathol Lab Med* November 1990; 114(11): 1101-1104.