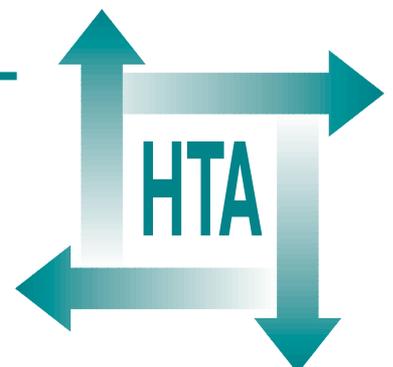# When and how to assess fast-changing technologies: a comparative study of medical applications of four generic technologies

G Mowatt
DJ Bower
JA Brebner
JA Cairns
AM Grant
L McKee

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

HTA

How to obtain copies of this and other HTA Programme reports.
An electronic version of this publication, in Adobe Acrobat format, is available for downloading free of charge for personal use from the HTA website (http://www.ncchta.org). A fully searchable CD-ROM is also available (see below).

Printed copies of HTA monographs cost £20 each (post and packing free in the UK) to both public **and** private sector purchasers from our Despatch Agents, York Publishing Services.

Non-UK purchasers will have to pay a small fee for post and packing. For European countries the cost is £2 per monograph and for the rest of the world £3 per monograph.

You can order HTA monographs from our Despatch Agents, York Publishing Services by:

– fax (with **credit card** or **official purchase order**)
– post (with **credit card** or **official purchase order** or **cheque**)
– phone during office hours (**credit card** only).

Additionally the HTA website allows you **either** to pay securely by credit card **or** to print out your order and then post or fax it.

Contact details are as follows:
York Publishing Services   Email: ncchta@yps-publishing.co.uk
PO Box 642   Tel: 0870 1616662
YORK YO31 7WX   Fax: 0870 1616663
UK   Fax from outside the UK: +44 1904 430868

NHS and public libraries can subscribe at a very reduced cost of £100 for each volume (normally comprising 30–40 titles). The commercial subscription rate is £300 per volume. Please contact York Publishing Services at the address above. Subscriptions can only be purchased for the current or forthcoming volume.

Payment methods

Paying by cheque
If you pay by cheque, the cheque must be in **pounds sterling,** made payable to *York Publishing Distribution* and drawn on a bank with a UK address.

Paying by credit card
The following cards are accepted by phone, fax, post or via the website ordering pages: Delta, Eurocard, Mastercard, Solo, Switch and Visa. We advise against sending credit card details in a plain email.

Paying by official purchase order
You can post or fax these, but they must be from public bodies (i.e. NHS or universities) within the UK. We cannot at present accept purchase orders from commercial companies or from outside the UK.

How do I get a copy of *HTA on CD*?

Please use the form on the HTA website (www.ncchta.org/htacd.htm). Or contact York Publishing Services (see contact details above) by email, post, fax or phone. *HTA on CD* is currently free of charge worldwide.

The website also provides information about the HTA Programme and lists the membership of the various committees.

When and how to assess fast-changing technologies: a comparative study of medical applications of four generic technologies

G Mowatt[1]   JA Cairns[4]
DJ Bower[2]   AM Grant[5]
JA Brebner[3]   L McKee[1]

[1] Department of Management Studies, University of Aberdeen,
    Aberdeen AB24 3QY
[2] Formerly Department of Management Studies, University of Aberdeen,
    Aberdeen AB24 3QY
[3] Department of General Practice, University of Aberdeen,
    Foresterhill, Aberdeen AB25 2AY
[4] Health Economics Research Unit, University of Aberdeen,
    Department of Public Health, Foresterhill, Aberdeen AB25 2ZD
[5] Health Services Research Unit, University of Aberdeen,
    Department of Public Health, Foresterhill, Aberdeen AB25 2ZD

NHS R&D HTA Programme

T he overall aim of the NHS R&D Health Technology Assessment (HTA) programme is to ensure that high-quality research information on the costs, effectiveness and broader impact of health technologies is produced in the most efficient way for those who use, manage and work in the NHS. Research is undertaken in those areas where the evidence will lead to the greatest benefits to patients, either through improved patient outcomes or the most efficient use of NHS resources.

The Standing Group on Health Technology advises on national priorities for health technology assessment. Six advisory panels assist the Standing Group in identifying and prioritising projects. These priorities are then considered by the HTA Commissioning Board supported by the National Coordinating Centre for HTA (NCCHTA).

This report is one of a series covering acute care, diagnostics and imaging, methodology, pharmaceuticals, population screening, and primary and community care. It was identified as a priority by the Methodology Panel (see inside back cover).

The views expressed in this publication are those of the authors and not necessarily those of the Standing Group, the Commissioning Board, the Panel members or the Department of Health.

Contents

Glossary and list of abbreviations

Technical terms and abbreviations are used throughout this report. The meaning is usually clear from the context but a glossary is provided for the non-specialist reader. In some cases usage differs in the literature but the term has a constant meaning throughout this review.

Glossary

The following definitions of different types of trials and related terminology. The definitions are taken from the glossary of the Cochrane Collaboration Handbook, version 3, compiled December 1996.

Case series An uncontrolled observational study involving an intervention and outcome for more than one person.

Case study An uncontrolled observational study involving an intervention and outcome for a single person.

Case–control study A study that starts with identification of people with the disease or outcome of interest (cases) and a suitable control group without the disease or outcome. The relationship of an attribute (intervention, exposure or risk factor) to the outcome of interest is examined by comparing the frequency or level of the attribute in the cases and controls. For example, to determine whether thalidomide caused birth defects a group of children with birth defects (cases) could be compared to a group of children without birth defects (controls). The groups would then be compared with respect to the proportion exposed to thalidomide through their mothers taking the tablets. Case–control studies are sometimes described as being retrospective as they are always performed looking back in time.

Clinical trial A trial that tests out a drug or other intervention to assess its effectiveness and safety. This general term encompasses randomised controlled trials and controlled clinical trials.

Cohort study An observational study in which a defined group of people (the cohort) is followed over time and outcomes are compared in subsets of the cohort who were exposed or not exposed, or exposed at different levels, to an intervention or other factor of interest. cohorts can be assembled in the present and followed into the future (a 'concurrent cohort study'), or identified from past records and followed from that time up to the present (a 'historical cohort study'). Because random allocation is not used, matching or statistical adjustment must be used to ensure that the comparison groups are as similar as possible.

Controlled clinical trial Refers to a study that compares one or more intervention groups to one or more comparison (control) groups. Whilst not all controlled studies are randomised, all randomised trials are controlled.

Cross-sectional study A study that examines the relationship between diseases (or other health related characteristics) and other variables of interest as they exist in a defined population at one particular time. The temporal sequence of cause and effect cannot necessarily be determined in a cross-sectional study.

Cross-over trial A type of clinical trial comparing two or more interventions in which the participants, upon completion of the course of one treatment are switched to another. For example, for a comparison of treatments A and B, half the participants are randomly allocated to receive them in the order A, B and half to receive them in the order B, A. A problem with this design is that the effects of the first treatment may carry over into the period when the second is given.

Observational study A study in which nature is allowed to take its course. Changes or

continued

differences in one characteristic (e.g. whether or not people received the intervention of interest) are studied in relation to changes or differences in other(s) (e.g. whether or not they died), without the intervention of the investigator. There is a greater risk of selection bias than in experimental studies.

Phase I studies The first stage in testing a new drug in humans. Usually performed on healthy volunteers without a comparison group.

Phase II studies Second stage in testing a new drug in humans. Often performed on healthy volunteers. These are sometimes randomised controlled trials.

Phase III studies Studies that are a full-scale evaluation of treatment. After a drug has been shown to be reasonably effective, it is essential to compare it with the current standard treatments for the same condition. Phase III studies are often randomised controlled trials.

Phase IV studies Studies that are concerned with post-marketing surveillance. They are often promotional exercises aimed at bringing a new drug to the attention of a large number of clinicians, and may be of limited scientific value.

Prospective study In evaluations of the effects of healthcare interventions, a study in which people are divided into groups that are exposed or not exposed to the intervention(s) of interest before the outcomes have occurred. Randomised controlled trials are always prospective studies and case control studies never are. Concurrent cohort studies are prospective studies, whereas historical cohort studies are not, although in epidemiology a prospective study is sometimes used as a synonym for cohort study.

Quasi-random allocation A method of allocating participants to different forms of care that is not truly random; for example allocation by date of birth, day of the week, medical record number, month of the year, or the order in which participants are included in the study (e.g. alternation).

Quasi-randomised trial A trial using a quasi-random method of allocating participants to different forms of care. There is a greater risk of selection bias in quasi-random trials where allocation is not adequately concealed compared with randomised controlled trials with adequate allocation concealment.

Random allocation A method that uses the play of chance to assign participants to comparison groups in a trial, for example, by using a random numbers table or a computer-generated random sequence. Random allocation implies that each individual or unit being entered into a trial has the same chance of receiving each of the possible interventions. It also implies that the probability that an individual will receive a particular intervention is independent of the probability that any other individual will receive the same intervention.

Randomised controlled trial An experiment in which investigators randomly allocate eligible people into (e.g. treatment and control) groups to receive or not to receive one or more interventions that are being compared. The results are assessed by comparing outcomes in the treatment and control groups.

Retrospective study A study in which the outcomes have occurred to the participants before the study commenced. Case–control studies are always retrospective, cohort studies sometimes are, randomised controlled trials never are.

Selection bias In assessments of the validity of studies of healthcare interventions, selection bias refers to systematic differences between comparison groups in prognosis or responsiveness to treatment. Random allocation with adequate concealment of allocation protects against selection bias. Other means of selecting who receives the intervention of interest, particularly leaving it up to the providers and recipients of care, are more prone to bias because decisions about care can be related to

prognosis and responsiveness to treatment.

1. Selection bias is sometimes used to describe a systematic error in reviews due to how studies are selected for inclusion. Publication bias is an example of this type of selection bias.

2. Selection bias, confusingly, is also sometimes used to describe a systematic difference in characteristics between those who are selected for study and those who are not. This affects the generalisability (external validity) of a study but not its (internal) validity.

List of abbreviations

ACOST	Advisory Council on Science and Technology	HTA	health technology assessment
CFTR	cystic fibrosis transmembrane conductance regulator	MAS	minimal access surgery
		MIS	minimal invasive surgery
CI	confidence interval	QALY	quality adjusted life-year
CODEC	coder–decoder	RAC	Recombinant DNA Advisory Committee
CONSORT	Consolidated Standards of Reporting Trials	RCT	randomised controlled trial
CVS	chorionic villus sampling	SERNIP	Safety and Efficacy Register of New Interventional Procedures of the Medical Royal Colleges
FDA	US Food and Drug Administration		
GP	general practitioner		

Executive summary

Objectives

- To try to identify the optimal time at which to start assessing new and fast-evolving health technologies.
- To provide insight into factors influencing the timing of assessments and the choice of methods for assessing new and fast-changing technologies.

How the research was conducted

A series of literature reviews were undertaken covering the general principles involved in the timing of health technology assessments (HTAs). Additionally, the reported assessments of laparoscopic cholecystectomy, chorionic villus sampling (CVS), teleradiology, teledermatology, genetic screening for predisposition to breast cancer, and gene therapy for cystic fibrosis were reviewed to try to identify the factors that influenced the timing of these assessments. Key individuals in each field were also interviewed. The selected technologies allowed comparison between those that were new and evolving and those that were relatively well-established.

A bibliometric study of publication trends was also undertaken to see whether these trends would suggest points in the development of a technology that could be used as indicators that assessment should be started.

Research findings

Timing

The precise point at which assessment should start was not identified but the bibliometric study suggested that extending this approach might give useful results.

For all health technologies, more regular reporting of outcomes and side-effects should be encouraged during the period after initial assessment and, where the technology is fast-changing, reassessment should take place from time to time. The precise intervals were not identified and the problem remains of deciding when a technology has changed enough to warrant reassessment.

Factors influencing timing

Published reports of assessments did not generally specify the reasons for their timing, but a number of factors appear to have influenced the timing of those assessments, directly or indirectly.

Product champions and opinion leaders pioneer the introduction of new technologies into clinical practice, and their reports may lead to the rapid diffusion of such technologies before they have been adequately evaluated, as was the case with laparoscopic cholecystectomy; this diffusion may limit the methods of evaluation that can then be used. It is therefore important to assess new health technologies before diffusion takes place.

The extent to which **regulatory control** is imposed on the introduction of new health technologies can also influence the timing of assessments. Such controls might have helped to restrict the diffusion of laparoscopic cholecystectomy, making a large and widely generalisable randomised controlled trial (RCT) feasible.

The **source and availability of funding** for studies may influence the nature and timing of trials. Many telemedicine evaluations were funded by commercial telecommunications organisations and were thus restricted in their timing (and biased towards the technological aspects of the applications) by the availability of funds.

Media coverage undoubtedly has an influence although this influence is not always predictable; it may generate 'favourable' publicity about new health technologies, which can lead to immediate demands for the new technique, as was the case with laparoscopic cholecystectomy with its apparent benefits. Thus assessments should be made before media coverage exerts popular pressure on purchasers to adopt the technology and dissuades patients from participating in RCTs (because of fear they may be randomised to the standard treatment as occurred in a US trial of CVS). Innovators should also be cautious in the claims that they make to the media.

Clinical uncertainty or equilibrium also affects the timing of assessments. During the period when clinicians have no preference between the

treatment options to be compared, they may be prepared to ask patients to participate in trials; however, once clinicians come to prefer either the standard or the alternative treatment, they may feel ethically obliged to provide only the treatment that they believe to be the best. This argument was given as a reason for abandoning a proposed RCT of laparoscopic versus open cholecystectomy. The counter-argument is that randomisation is a hedging, risk-minimisation strategy when the true risks and benefits are not known.

The existence of the **clinical learning curve** also influences the timing. Assessments made before clinicians have acquired enough skill in the new procedure may produce misleading findings on benefits and costs. Assessments may need to be postponed until clinicians have reached an appropriate point on the learning curve but this can usually only be recognised retrospectively, by which time clinicians may no longer be prepared to randomise patients.

The fact that the development of some health technologies is **technology-driven** or **commercially-driven**, rather than needs-based, affects the timing of assessments to the extent that advances in the technology, in conjunction with a **reduction in costs**, have largely determined the timing of assessments (e.g. teleradiology).

Assessment methods

HTA has traditionally focused on clinical outcomes but there are now demands for a wider range of criteria including social and ethical impact, effect on patterns of healthcare demand, cost-effectiveness and other issues.

The reviews of the various applications indicate that HTA can never be perfect but that best practice uses a number of methods of assessment, rigorously applied and reported, to achieve the most satisfactory outcomes for patients. Rare side-effects are often only detected after extensive use, and new problems can arise because of the different ethical and cultural concerns of different patient groups. New patterns of demand, created by the availability of new techniques, can invalidate economic studies.

The problems of assessing fast-changing technologies are similar to those of assessing stable technologies but are likely to arise more often

during the development phase of a technology. Regulation is restricting the genetic technologies to research use or controlling their diffusion until assessment gives satisfactory outcomes. Telemedicine, however, is only beginning to be assessed on a limited basis and there are no controls on adoption. Thus, approaches to assessment are more a function of perceived risk than of rate of change.

Both stable and fast-evolving technologies lack a framework of standard guidelines and incentives to ensure that users assess unregulated or lightly regulated health technologies in an approved and consistent way and report the results. In addition, guidelines are needed to ensure that the decision about when a procedure has changed enough to be regarded as new is clearer and less subjective.

Conclusions

- Assessment should be initiated early, using a variety of complementary assessment approaches.
- Methods of assessment and reporting should be more standardised from the earliest stages, to improve the usefulness and comparability of data.
- Resource issues should be incorporated into assessments from an early stage.
- All technologies should not be dealt with in the same way – they should be assigned to categories for which appropriate common triggers can be identified.
- Trials should be randomised from the outset.
- Assessment should be an iterative process.
- Citations and publication trends may be useful for identifying triggers.

Research recommendations

- Bibliometric studies involving a larger number of established technologies should be undertaken to detect whether there is a sufficiently consistent pattern to the publication trends of new and fast-changing health technologies to allow identification of a 'critical point' at which assessment should be recommended.
- Guidelines for the study and interpretation of different types of health technologies should be developed to facilitate assessment decisions.

Chapter 1

Introduction

Rationale for the study

There is currently no accepted formula for the timing of health technology assessment (HTA) of new and fast-evolving technologies. By their nature, they are often difficult to evaluate by accepted HTA methods. However, there are examples of unevaluated new technologies that have diffused widely and have subsequently been shown to be ineffective or even harmful (Stocking, 1985; ACOST, 1993). Consequently both the safety and the cost-effectiveness of service delivery would be improved if useful criteria were determined for guiding the timing and nature of assessment, and for deciding when new developments should become routine services.

An additional problem is posed by the growing diversity of professionals who make decisions about the adoption of new technologies. Depending on a number of factors, such as the amount of controversy surrounding the generic technology (e.g. genetic manipulation), decision-making may be subjected to central monitoring or to a formal regulatory process to varying degrees of decentralised regulation or it may be unregulated. Although, in the case of ethical pharmaceuticals, regulations have been imposed in response to perceived physical harm to patients, routine systems have not yet been devised for the evaluation and control of many other types of interventions which also have obvious potential for harming patients. There is a heightened awareness of the need to evaluate new technologies and there may even be devolution of responsibility for commissioning, managing and implementing some evaluations, where appropriate (Culyer, 1994). It would, therefore, be particularly useful if criteria for timing assessments and for the assessments themselves were developed and translated into a series of formats suitable for a range of circumstances, including one for managers and others who lack research expertise.

The aim of the study reported here was to provide insights into factors influencing the timing and the choice of methods of assessment of new and fast-changing technologies.

Diffusion of unevaluated medical technologies

Previous research has shown that new medical technologies have sometimes spread rapidly in the absence of any evidence confirming their effectiveness (Stocking, 1985; 1991; Banta, 1993; ACOST, 1993). In spite of policy designed to induce a more evaluative culture in the NHS (Peckham, 1991a; b; 1993), there are now many technologies available for medical application that are likely to diffuse widely before evaluation. In addition to the general characteristics that influence rate of diffusion, such as cost, the presence of enthusiasts, lack of resistance, meeting perceived needs and ease of use (Rogers, 1983), a number of factors, specific to healthcare delivery services, that affect the speed of diffusion are listed below.

- *Non-medical evolution* occurs where the generic technologies are evolving rapidly in non-medical applications, for example, information technology and interactive video-conferencing. Developments from generic technologies designed for other purposes, then hastily adapted for medical use, can diffuse rapidly into medicine.
- *Powerful popular pressures* have been summed-up by the phrase 'the need to do something' (Stocking, 1985). Patient demand, intensified by media reports of availability of susceptibility tests, is creating a wider demand, for example, for testing for many conditions about which the benefits of knowledge are doubtful or for which, as in the case of breast cancer, after diagnosis of susceptibility there is not, as yet, any known management strategy that offers net benefits (see, for example, Bower, 1994). In addition, the availability of new technologies in the USA and other markets puts pressure on the NHS to adopt without evaluation.
- *Powerful commercial pressures* – while governments regulate corporate promotion of healthcare products to some degree, there is a widespread perception that this is not wholly effective. Since pharmaceutical companies spend, on average, 30% of revenues on marketing (see their Annual Reports), they clearly share this perception.

- *Lack of regulatory barriers and difficulty and/or expense of evaluation* – as in the case of telemedicine – are particular problems where human expertise and organisational arrangements can be expected to affect outcomes significantly (see, for example, Stocking, 1985).
- *Methodological limitations* make it difficult to find appropriate, replicable research/evaluation methods for healthcare technology innovation. This is indicated, for example, by reliance on mortality data and limited alternative outcome parameters.

The task of evaluation

Most medical technologies are complex (Comroe & Dripps, 1981): they bring together elements of 'hardware', such as instruments and drugs, with a 'software' component, the clinical knowledge and expertise gained through general training and training specific to the application (Rogers, 1983). For instance, telemedicine uses videoconferencing and other hardware but also has an important component of 'software' – the specialist knowledge and expertise of the clinician in interpreting whatever visual and auditory information he/she receives through the link. Where the knowledge/expertise component is significant and non-routine, evaluation is complicated by the difficulty of making comparisons.

The task of evaluation is easiest when the dominant elements are highly testable and routinised, and when regulation requires their testing in specified situations. However, the presence of extensive, expensive-to-satisfy regulations designed to ensure the safety and efficacy of drugs has created the situation in which only those drugs that can be evaluated conventionally (largely in randomised controlled trials (RCTs)) are developed by companies. In general, where regulation has been imposed, company interest focuses on products that can meet regulatory requirements.

While regulation can exert some control over diffusion, it would be undesirable to impose new regulations in any situation in which this might lead to the neglect of important possibilities for improving the cost-effectiveness of health-care. Regulation has undoubtedly led pharmaceutical companies to neglect many areas in which their technology might offer effective therapies, leading, in the USA for example, to special legislation to encourage such developments, viz. *The Orphan Drug Act, 1983* (96Stat.2049, 21USC301). Consequently it might be preferable to devise useful and cost-effective assessment methods that, unlike costly rigid statutory regulation, do not suppress innovation.

The increasing decentralisation of the NHS also raises novel issues for the task of evaluation, since some responsibility for decision-making about research and development and assessment, particularly for unregulated technologies, is likely to be borne by hospital managers.

Timing and nature of assessments

Following Rogers' (1983) model, Stocking (1988) has pointed out that an important time to assess a new technology is "at the point when opinion-leaders become interested in it", which is at a very early stage, often before clinical trial data are available. After this, the involvement of a multiplicity of powerful constituencies restricts the possibility of controlled evaluation before diffusion. The question of how to identify this point then arises. The date and rate of increase in the frequency of conference research papers authored by opinion leaders on a technology might be possible indicators of the birth of interest and growth of commitment. However, where a technology is fast-changing, this raises questions as to what type of useful assessment can be carried out cost-effectively at this stage.

The point at which a research application is allowed to move to routine use is also important, with the consequent implications for capital and running costs, training and so on. These implications may be difficult to predict accurately: although some discussion of the direct cost implications of adoption of minimally invasive surgery (MIS) was included in a special issue of *Health Policy* (Banta, 1993) devoted to evaluations of MIS applications, it was also noted that the changes in philosophy driven by this technology would probably lead to much greater changes in the pattern of demand for hospital buildings and training (Wickham; in Banta, 1993).

These are points at which decisions with major cost and/or control implications are made. They also reflect the transition to a period of greater uncertainty about the utility of a technology and about the extent of long-term service implications.

Another issue that might influence the timing (and nature) of assessment is popular acceptability, particularly where there are ethical concerns, as in the case of genetic screening. For example, the Nuffield report (Nuffield Council on Bioethics, 1993) is an assessment of current and probable future possibilities for genetic screening, focusing on which technological trajectories are or are not acceptable in the ethical climate of the UK. The assessment does not address issues of cost but underlines incidentally the difficulties of defining

'efficacy' in the context of interventions to which different cultural groups attach very different values.

In the case of rapidly-evolving technologies, there are also questions about the probability of new medical applications emerging and where the trajectory of change is leading the technology. Is it becoming significantly cheaper each year, thus affecting cost estimates? Is it becoming more precise and controllable, affecting efficacy? Have unexpected adverse effects been reported that call into question the general viability of the technique? Is there growing or declining scientific interest, signalling a change in rate of evolution?

Consequently both the nature and timing of assessments of new and fast-evolving technologies involve a number of additional issues when compared with HTAs of routinised interventions.

The research questions

The research questions addressed by the present study are listed below.

1. What general principles have been reported as guiding the timing of HTAs in the past? Are they adequate for new, fast-changing technologies?
2. What gaps does a literature review reveal in the reporting of the principles that guide the timing of HTAs?
3. What further insights can be gleaned by interviewing key personnel associated with the assessments reviewed?
4. How and when have reported assessments of the specific applications of the four generic technologies listed in *Table 1* been carried out? Have opportunities for practical and desirable assessments been lost? Are there lessons to be learnt from the applications that have now

reached a relatively stable state, which could be applied to the fast-evolving applications? To what degree are assessments comparable at different points in the evolution of a fast-evolving technology?
5. Is it possible to draw inferences about trends in the development/diffusion of technologies from collective characteristics of reported assessments, that can be used to decide on the timing of assessments? For example, can trends in the rate of publication or citation in the scientific literature, or trends in reporting in the popular media, be used as criteria for **timing** an evaluation? Can such trends be used to pinpoint the right moment to issue formal reminders about the need for evaluation, and the way in which it should be implemented, in order to pre-empt the rapid diffusion of an unevaluated technique?
6. Following on from question 5, is there inform-ation within the literature reviewed that can be used to derive useful general guidelines and protocols for assessment? How much of a problem is posed by issues specific to the technology and application?
7. Does the information gleaned from the reviews yield insights as to when 'new' technologies should become part of routine practice?

Methodological approach

Three complementary methods were used to address the research questions. First, a series of systematic literature reviews were carried out focus-ing specifically on the timing and method of assess-ment (as described below). Second, since the con-ventional format of papers in the academic liter-ature rarely includes explicit indicators about the choice of timing, key individuals associated with the reviewed health technologies were interviewed to elucidate this further and to ascertain whether factors other than those reported influenced the choice of methods. Third, a bibliometric study of publication trends associated with the health tech-nologies reviewed was undertaken, to explore the possibility that particular events might produce char-acteristic patterns in the trends, thus providing an indicator which could be used to signal the need to initiate an assessment. A number of modifications to the protocol as originally proposed were introduced to optimise the overall quality of the study, and these are discussed in the Conclusions (chapter 9).

Systematic reviews
A series of systematic literature reviews were carried out:

TABLE 1 *Applications reviewed and their generic technologies*

Generic technology	Application
Karyotyping	CVS
MIS	Cholecystectomy
Genetic manipulation	1. Gene therapy (cystic fibrosis) 2. Diagnosis of genetic susceptibility to breast cancer
Videoconferencing (telemedicine)	1. Radiology 2. Dermatology

(i) of studies specifically addressing the question of the timing of assessments of technology

(ii) of relevant literature reporting the assessment of six medical applications of four generic technologies. The applications included two relatively well-evaluated and technically fairly stable applications, and four new and rapidly evolving applications *(Table 1)*.

Karyotyping of human tissue samples has been used for diagnostic purposes for more than 25 years. Chorionic villus sampling (CVS) is an application of this technology which was developed in the early 1980s to diagnose foetal karyotypic abnormalities (mainly trisomy 21) at an earlier stage and with less maternal discomfort than other available techniques. It was extensively evaluated and was found to generate serious side-effects at a level which, although low, was higher than for alternative methods.

MIS involves the use of visualisation of the site of surgery and remote control of surgical instruments by electronic techniques through a small aperture in the patient's body. It is a recent modification of remote manipulation technology long used in non-medical contexts, for example, to disassemble radioactive fuel rods. Cholecystectomy by MIS is a recent (1990s) but extensively evaluated application (Cuschieri A: personal communication, 1996).

Genetic manipulation technology involves the production of defined nucleic acid sequences with known functional properties. Medical applications under development, such as those described below, are presently causing much controversy.

1. Cystic fibrosis gene therapy, currently undergoing tightly regulated development at several sites in the USA, France and the UK, is an invasive application in which normal copies of the defective cystic fibrosis gene are introduced into the tissues of affected individuals in order to stimulate local synthesis of the missing essential protein.

2. Diagnosis of genetic susceptibility to breast cancer uses several recently isolated sequences known to be carried by individuals with heightened susceptibility to breast and other cancers. Accuracy of diagnosis is increasing rapidly but there is, as yet, no management strategy which is known to offer net benefits. A major EU-funded comparative study of its development at several sites in Europe, coordinated by Professor Neva Haites (Aberdeen University), has recently been initiated.

Videoconferencing (telemedicine) involves the use of commercial videoconferencing hardware and ISDN (Integrated Services Digital Network) telephone links. Although pioneered over 30 years ago, hardware of a quality adequate for routine use has only recently become commercially available. The first entry under 'telemedicine' in Medline as a textword term occurred in 1974, and as a MeSH term in 1991. Emerging applications are, thus far, largely unevaluated. Radiological and dermatological applications are forecast to be two of the most likely to be widely used.

Criteria for selection

The generic technologies and applications have been selected to allow meaningful comparisons between an adequate range of technologies, of various ages and at various stages of evaluation, so as to gain generally applicable insights into the problems of how and when to assess new and fast-changing technologies. For instance, they permit pertinent comparison between applications that are unregulated and those that are highly regulated *(Table 2)*, between those with a low media profile and those with a higher profile, between commercially-driven and publicly-funded applications, and between diagnostic and therapeutic technologies. There may be some general principles that apply to all applications, but there may also be some that apply differentially to, for example, non-invasive versus invasive applications and some that are specific to the generic technology rather than to the medical application.

All four generic technologies have had non-medical applications for some time and, for some of them, it is not clear when precisely they could be regarded as 'emerging medical technologies'. For example, telemedicine has been used in clinical settings in a very limited way for 30 years,

TABLE 2 *Level of regulation and invasiveness of reviewed applications*

Application	Level of regulation	Invasiveness
CVS	Local	High
Laparoscopic cholecystectomy	Local	High
Gene therapy (cystic fibrosis)	Central	High
Diagnosis of genetic susceptibility to breast cancer	Local	Low
Teleradiology	None	Low
Teledermatology	None	Low

while medical applications of MIS are more recent. However, the infrastructure and hardware that make telemedicine feasible for widespread use have only recently become affordable, while some applications of MIS, although quite recent, have been extensively evaluated, for example in gynaecology, and the technology involved has become quite stable.

Consequently, the main distinction is between, on the one hand, relatively stable and well-evaluated technologies and, on the other, fast-changing technologies which in some cases also pose novel problems from the point of view of evaluation. The generic technologies fall into two technologically different categories, 'tele-' and 'chromosome/gene-' based technologies, each of which offers useful paired comparisons between stable, well-evaluated and fast-changing technologies. MIS offers close parallels with videoconferencing, in terms of the nature of the technology, its non-medical origins, and the novel problems it poses for effective use: the visual simulation of the target site poses comparable challenges for the human brain in the input, processing and utilisation of the information and in remote control of actions, and the non-medical applications contribute to the evolution of the technology. Karyotyping and genetic manipulation use classical and novel genetic technologies in which most development has been undertaken for medical applications.

The final criterion for selection was that, to undertake the project, the authors should have the necessary knowledge and technical expertise in relevant aspects of the generic and applications technologies, together with access to other key research groups.

Objective of the reviews

The objective of the reviews was to analyse the nature and quantity of assessments of the above applications and to clarify factors that determined the timing of such assessments. It appeared probable at the outset that these literature reviews would not, in themselves, provide all the information needed to answer the questions of when and how to evaluate any new or fast-changing medical technology. However, any inferences that could be drawn from these reviews on:

- the utility, cost and timing of specific types of assessment for NHS decision-making on the funding of further research and development
- patient and public acceptability

- when to permit an intervention to move from 'research' to 'routine' status
- long-term resource implications

would help to strengthen the basis of research and development decision-making in the NHS. It was not within the scope of the reviews to evaluate the quality of the assessments described in the papers reviewed, since useful information about timing decisions and choice of methods could be gleaned even from methodologically flawed studies. However, this does not mean that the quality of a study and its timing are unrelated.

Because of the nature of the topic, many publications, especially on the general principles of HTA, represent the views of the authors rather than relating to the outcomes of clinical trials. The extent to which the six medical applications have been evaluated by, for example, RCTs, has been variable because of factors such as the perceived degree of risk. If the references cited relate to clinical trials or specific types of clinical trials, such as RCTs or observational studies, then this is stated. A complete listing of all references cited, not just clinical trials, is given in the reference list, and the lists of references for each specific medical application contain all those cited in the relevant chapters, not just the clinical trials.

Within the time and resource limitations of the study it was planned to locate where broadly applicable general rules could be generated (if at all). The sample thus constituted a useful combination of generic and applications technologies for an initial investigation of the wider questions of the timing and nature of assessments. It was judged that they would give an indication of the degree to which issues are specific or general, and also of the domain of the assessment issues of new and fast-changing technologies and whether these were specific or general.

The reviews would also reveal some of the information gaps that could be filled by other methods, such as interviews with key individuals, and also indicate when opportunities to undertake assessments had been grasped and when such opportunities had been missed. Thus, they would assist in the definition of future research requirements.

Review strategy
The reviews of the literature on the technologies of the six applications were performed using a systematic search strategy. The period of search extended from the point at which the idea for a particular medical application was reported (before the first

clinical trial) until November 1996. The target literature included that which predicted the specific clinical applications of the generic technology before first use, as well as subsequent reports of clinical trials and assessments relating to the applications. This strategy was designed to address both the question of the practicability of timing evaluation before uncontrolled diffusion and the question of the **nature** of any evaluation that had been undertaken.

Reviews started from the interpretation of 'what can be assessed', which has been applied to the older applications, and systematically determined the utility of different types of assessment (technical feasibility studies, RCTs, etc.) where possible. Information on medically important modifications of technique (cost, ease of use, precision, novel instrumentation, etc.), technological trends as they affect medical applications, patient and public acceptability, commitment of substantial public or private funds (including non-UK), and known barriers to routinisation, as well as clinical outcomes, were sought which might be used for assessment. Economic aspects of published HTAs of the applications were discussed when available.

Relevant literature
Relevant literature may be classified into three groups.

1. The primary medical research literature was the main source but reviews of trials (i.e. secondary level evaluation) were also used.
   - Relevant publications were identified using the electronic databases, Medline and Embase, with an empirically-determined optimal combination of MeSH terms. The databases of the Centre of Reviews and Dissemination, York, were consulted. Science Citation Index (Scisearch) was also used both to identify key articles in terms of citation rate and to identify the articles that were citing them.
   - Authors of evaluations and others known to be conducting evaluations or to have an interest in the field were contacted for information on other relevant publications and to ascertain further details of their evaluations, including criteria for timing.
   - The bibliographies of studies identified by electronic searching and bibliographies that had already been generated were searched.
   - Other languages – although non-English literature could not be searched with the same thoroughness, the aim was to limit the risk of bias by reviewing relevant publications in a limited number of major journals and newspapers published in French and German. Translation was carried out in the UK.
2. Reports of significant study groups/working parties such as the Advisory Council on Science and Technology (ACOST) and the Nuffield Council on Bioethics, were hand-searched for relevant commentaries on timing and types of assessment. Conference proceedings were also scanned for relevant articles.
3. Press reports were identified using the electronic database Promt (which contains full text articles from worldwide business periodicals, including the main daily national and regional newspapers).

Bibliographic database
Articles to be included in each review were organised into a series of databases using the software package, Reference Manager®. Each article was summarised as an abstract which included a general summary of its contents and relevance to the study.

Interviews with key individuals
When the reviews were under way, it became apparent that much of the required information was missing. The systematic reviews yielded little insight into the timing of the assessments reported and were not always explicit about the choice of methods. In the case of CVS, one author (AMG) had been directly involved in the assessment and was able to provide important background information on timing and the selection of methods. Key individuals for the other applications were interviewed to provide corresponding insights; this was a speedy and effective way of obtaining relevant information that is not normally reported in academic papers.

The key individuals interviewed were:

- Professor A Cuschieri (Dundee University), who has been extensively involved in the development and evaluation of MIS
- Professor Neva Haites (Aberdeen University), who is responsible for genetic screening for breast cancer and is coordinating a European survey of this
- Professor David Porteous (MRC Human Genetics Unit, Edinburgh), who is coordinating multicentre clinical trials of cystic fibrosis gene therapy, funded by the MRC
- Dr Ross Maclean (HSRU, Aberdeen University), who has been extensively involved in the development and evaluation of telemedicine applications.

Bibliometric study

Literature-based innovation output indicators (Edwards & Gordon, 1984) are a useful addition to the range of indicators of innovation activity. The pros and cons of their use has been discussed in some depth by Coombs and colleagues (1996). The advent of many large computerised databases of publications of all types, from papers in academic journals to daily newspapers, has made it relatively easy to use numbers, types, contents and other characteristics of publications as indicators of reported activities. These bibliometric methods have been used by a number of researchers to analyse the development and composition of research and development communities (see, for example, Garud & Rappa, 1992; Rappa *et al,* 1992; Clarysse *et al,* 1996). Using a set of search terms, the method was used in the study reported here to assess the size of the research and development communities associated with the applications reviewed (from numbers of publications per year) and the nature of the coverage (from the content of papers). Content of publications was also related to crude trends in rate of publication. The rationale was that those points at which significant changes in interest occurred might be indicated by changes in publication trends. The 'contribution rate' of scientists (publication rate within a given literature) has been used in other industries as a measure of the rate of emergence of a new technology (DeMeyer *et al,* 1994). The relationship (if any) between these measures and assessment decisions was considered.

Initial assumptions

The starting assumption of the bibliometric strategy was that numbers of publications and trends in publication about an application were related **in some way** to the rate of evolution of a particular medical technology. The relationship was not expected to be a simple correlation but it was predicted that, in periods when the number of publications about an application was relatively high, its evolution would be faster than in periods when this number was lower.

The second assumption was that if the number of publications about an application was increasing year by year, over a period, the rate of evolution would be rising. Conversely, when the number fell, the rate of evolution would be slowing and the technology might be reaching the status of a 'dominant design' which would be likely to remain stable for a prolonged period (Utterback, 1994). These trends for different categories could be directly plotted against one another. The assumptions were tested by comparing the publication trends for the relatively stable technologies with those of the fast-evolving technologies. The relationship of initiation of assessment and reports of outcomes of assessments to changes in trends was also evaluated, to give some insight into whether publication trends could in themselves yield useful information to guide the timing of assessments.

Literature to be subjected to bibliometric analysis

Five types of literature were reviewed on a year-by-year basis, starting from the present and working back to a 'first idea' – the point at which a technically well-informed writer first suggested the application using that generic technology **in print**. The types of literature selected to give information about the impact of 'technology push', clinical interest and popular/patient influence on the rate and direction of evolution of these technologies were as follows:

* technically well-informed literature on 'ideas for medical applications' (including conference papers, general reviews of the state of the technology in scientific, engineering and medical journals, and reports commissioned by official (government) bodies)
* scientific literature on clinical applications
* scientific literature on assessments of applications
* popular literature on 'ideas for medical applications' (including six major English language newspapers published in the USA or the UK, one French and one German newspaper)
* popular literature on clinical applications (reports of clinical trials in the same eight newspapers).

If an identifiable and similar pattern of reporting of particular events preceded the first clinical trials of applications, such a pattern could constitute an indicator and used to trigger a process of evaluation.

The authors' expertise

The interpretation of findings was supported by the expertise of the research team Their combined specialist expertise extended across the wide range demanded by the project – expertise in the technologies as they were being applied in medicine, in NHS organisation and management, in the evaluation of efficacy, effectiveness and health economic aspects of medical technology, and in preparing systematic reviews. Details of the expertise of individual authors is presented in Appendix 5.

Presentation

The results of the study are presented as follows.

1. A systematic review of the academic literature on the timing and nature of assessments – chapter 2.

2. Publication trends for a given application in the literature – chapter 3.

3. Systematic reviews of the six applications considered – chapters 4–8.

4. The qualitative and quantitative changes in publication that were found were correlated with reported start dates of clinical trials, and other key events that might be expected to impact on the pattern of development, including media reporting – chapters 4–8, as appropriate, and chapter 9.

5. Discussion and recommendations for future research – chapter 9.

6. List of professional conferences, seminars and other meetings – Appendix 2.

Chapter 2

When should a health technology assessment be initiated?

Summary

There is currently no generally accepted formula for the optimal timing of HTAs. Some of the relevant issues are presented in this chapter and the existing literature on timing of HTAs is reviewed. Literature that specifically addresses these issues is limited. There is a consensus that HTAs should be initiated at an early stage of the development of a new health technology and repeated during the life cycle of the technology; however, the problems of reliably identifying such new technologies at an early stage in their development and of deciding on a detectable critical point for starting evaluation are not resolved. It is proposed that a system of categorisation and prioritisation of health technologies should be developed so that decisions can be made about when a strongly precautionary approach needed and how the limited resources available for HTA can be optimally deployed.

Background

Health technologies have been defined as "any method used by health professionals to promote health, prevent and treat disease, and improve rehabilitation and long-term care" (Standing Group on Health Technology, 1994). Such methods include both drugs and devices considered as technologies in their own right, together with medical and surgical procedures which may themselves involve the use of drugs and devices. Health technologies also embrace the higher level organisational systems that are vehicles for these procedures.

A major objective of HTA is to provide patients and clinicians with information on patient care alternatives, and to provide policy decision-makers and healthcare managers with information on alternatives (Donaldson & Sox, 1992). As a consequence, HTA involves the evaluation of the benefits and costs (clinical, social, economic and system-wide) of transferring the technology into clinical practice (Russell, 1996). Health technologies at each of the three levels of complexity described above should, in theory at least, be evaluated in

these ways. Ideally, HTA should be thought of as a comprehensive form of policy research that examines both short- and long-term consequences of the application of technologies (Luce & Brown, 1995).

In the case of drugs, there is a well-defined and fairly comprehensive regulatory route of HTA, which is refined, from time to time, when inadequacies are detected. For other health technologies, however, both the type of evaluation that is appropriate and the timing of such an evaluation have received much less attention. Most responsibility for the adoption of such technologies is delegated to individual clinicians and local research ethics committees, who must make decisions aided only by fragmented, often vague and incomplete recommendations and guidelines handed down by a variety of professional and governmental bodies, and without specific resource provision. This situation arises because of a lack of detailed central guidance, rather than because of an informed decision that this is the most appropriate way of dealing with the questions that arise.

The pattern of development and adoption of HTs

One of the reasons for the absence of detailed guidelines for assessment is the lack of agreement about how and when new technologies can and should be evaluated. According to Banta and Luce (1993), the lifecycle of a technology consists of five stages:

(i)   future – technology not yet developed
(ii) emerging – technology prior to adoption
(iii) new – technology in the phase of adoption
(iv) accepted – technology in general use
(v)   obsolete – technology that should be taken out of use.

Szczepura (1993) defined new technologies as those that had recently been introduced. The Department of Health (1995), in a document outlining a proposed Safety and Efficacy Register of New Interventional Procedures (SERNIP) by the Academy of Medical Royal Colleges, defined a new interventional procedure as "an invasive procedure which a clinician has read about, or has heard

about, or has piloted (following Local Ethics Research Committee approval), but for which either the safety or the efficacy of the intervention has not been established. It does not include minor modifications of existing procedures where the safety and efficacy are not in question."

Although technologies may evolve through the stages referred to above, the process is not necessarily a straightforward progression. The earlier view, that technologies, including health, had a simple and consistent pattern of development and diffusion, has been largely abandoned in favour of a much more complex model in which it is difficult to identify clear transition points. The simplistic, older pattern of the life cycle of a health technology consisted of its development, adoption, obsolescence and eventual abandonment (see, for example, Banta & Thacker, 1990). This life cycle followed a sigmoid or 'S'-shaped curve, in which the initial slow diffusion was followed by a phase of rapidly increasing adoption, which reached a plateau as saturation occurred (Feeny *et al,* 1986).

Gelijns and Rosenberg (1994) criticised this linear model, in which an idea for a new health technology moves in an orderly progression from the laboratory to animal models, to select human populations, and finally into routine practice. While this model does approximate to the highly constrained pattern of tightly regulated technologies such as ethical pharmaceuticals (i.e. basic research – applied research – targeted development – manufacture and marketing – adoption), it implies that technological innovation is much more systematic than it really is. In fact, most technologies develop in slightly different ways at a number of sites simultaneously. There is abundant evidence that their adoption and diffusion follow an irregular path, with different centres modifying them in idiosyncratic ways, with little concern for meaningful evaluation. Deber (1992) observed that, in general, technologies diffuse incrementally, rather than through an orderly process of assessment and that this results in pilot projects becoming established procedures despite the absence of evaluation.

Another criticism of the linear model made by Gelijns (1990) is that it does not take account of the fact that stages of the development process are influenced not only by research but also by the broader environment through market forces. She argues that technological development is an iterative process, in which both an underlying and evolving scientific and engineering knowledge base (technology-push) and market demand

(demand-pull) interact resulting in a particular pattern of innovation.

A further complication is the introduction into medical use of technologies already developed for other purposes by different industries. The adaptation of videoconferencing for telemedicine applications exemplifies this. In such cases, bodies undertaking clinical evaluation are unaware of the early stages and much of the subsequent development of the technology. Gelijns and Rosenberg (1994) argued that a high percentage of new medical devices have emerged, not from biomedical research, but through transfer of technologies developed elsewhere. This ability to evolve rapidly and independently of the medical context adds to the unpredictability and volatility of the emergence of new technologies.

Changes in the way in which health-care is purchased and delivered may also affect the pattern of the innovation process. Traditionally, physicians, acting as agents for their patients, were considered by the developers of new health technologies to be the principal users. In recent years, however, other groups – such as policy makers, hospital managers, patients, prospective patients (i.e. the general public) and regulators have begun to affect the demand for technology. These groups have an important influence on which new technologies will be accepted into practice and how they will be used (Bower, 1994). This, in turn, affects the rate and direction of subsequent research and development efforts (Gelijns & Rosenberg, 1994).

Development does not end with the adoption of an innovation. Adoption can be the beginning of an often prolonged process in which important redesigning takes place, incorporating the feedback of new information generated by users. Thus, a critical characteristic of unregulated innovation in medicine today is that new technologies retain a high degree of uncertainty long after their initial adoption (Gelijns & Rosenberg, 1994). This raises a difficult question – how do you evaluate a health technology which is changing so rapidly that by the time the results of an evaluation are available they may be irrelevant?

Resource constraints on evaluation
Another constraint on HTA is the inevitable expense associated with any formal evaluation. This is particularly problematic for fast-changing technologies where the results may be outdated by the time an evaluation is completed. Costs will vary with the nature and extent of the evaluation

process but are always incurred. The more extensive the evaluation, the greater the cost. At the high end of the cost spectrum comes the highly-regulated drug-development process. With estimates of $300 million as the cost of successfully developing one drug through to clinical use, mainly because of the cost of meeting regulatory requirements (DiMasi, 1991), there is some reluctance to impose central regulation on other types of health technology without definite justification.

However, it could be argued that the costs of adequate evaluation are a small price to pay if ineffective or harmful technologies are prevented from diffusing into routine practice. Gelijns (1990) argued that inconsistencies between the development of drugs, devices and procedures during the development process might have contributed to unnecessary healthcare costs, if allowance had been made for the fact that the least systematically evaluated technologies – surgical procedures – were also the most costly.

Public demand
Rigorous HTA also involves some limitation of the wider availability of a technology while it is under evaluation. In the absence of regulation, this limitation requires the agreement of both the public and the clinicians who might use the technology. Such agreement, in turn, requires acceptance by these parties that evaluation is necessary – and this is not always forthcoming. When hopes are raised that a new technology offers a major advance over current practice, political pressures to adopt an untested technology may develop. Such political pressures are brought to bear with conspicuous effect when backed by the power of the media.

The recognition of the complexity and irregularity of the innovation process, of concerns over the costs of ineffective treatments, and of the need for guidelines to help withstand pressures to adopt untested technologies, have created an imperative to seek convincing indicators of when and how health technologies should be assessed. The literature on the timing of HTAs is reviewed below, and the conclusions of the researchers are discussed.

Differences in the control of drugs and surgical procedures
Nearly 20 years ago Bunker and colleagues (1978) drew attention to the marked differences in the controls and regulations governing the introduction of new drugs compared with new surgical procedures. They pointed out that new drugs had to conform to strict, centrally-imposed regulations, which required rigorous testing in animals, according to strict experimental designs, followed by carefully controlled testing in humans, with appropriate protocols and follow-up observation. In contrast, new surgical procedures were usually introduced in an uncontrolled manner, free of any form of regulation. Sheldon and Faulkner (1996) argued that this has allowed a tidal wave of new healthcare technologies to diffuse through healthcare systems before proper evaluation has been undertaken to establish their safety, effectiveness, or return on investment.

Love (1975), however, argued that it was inappropriate to draw parallels between the development of drugs and surgical procedures. Important differences between drugs and operations included the element of manual skill required to perform surgery, and surgical techniques, unlike drugs, did not have chemical compositions, physical properties, or other qualities that could be measured precisely. He pointed out that the details of an operation typically evolved with experience, and the skill factor varied not only between surgeons but also in the development of individual surgeons as they became familiar with a new procedure. Assessment of safety and effectiveness could only be made when properly applied in the clinical setting.

Clinical procedures
Gelijns (1990) pointed out that in contrast to drugs or devices, no formal governmental regulatory system existed for the development and evaluation of clinical procedures. Their development had traditionally taken place in the context of the physician's clinical autonomy and the trust between physician and patient; evaluation of these procedures during development depended to a great extent on professional self-regulation. As a consequence, Gelijns argued, the potential safety, efficacy and effectiveness of many procedures had not been evaluated systematically during development.

The institutional structure within which development decision-making took place differed to some extent for devices, drugs, and surgical procedures, according to Gelijns (1990). The development of drugs and devices was largely sponsored by the pharmaceutical, biotechnology, and medical device industries, and took place both in these industries and in academic and governmental clinical research settings, where investigators evaluated the likelihood of benefits and risks in patients. Procedures, on the other hand,

were both technically developed and clinically evaluated by physicians in clinical practice (Gelijns, 1990).

From her study, Gelijns concluded that there were serious inconsistencies in the evaluation of drugs, devices and procedures during their development; these might contribute to shortcomings in the effectiveness and efficiency with which biomedical research findings and clinical theories were translated into useful clinical practice.

Methods of assessment

No attempt is made here to review the extensive literature on **methods** of HTA but, since the timing and choice of method of assessment are interrelated, a brief discussion of the types of assessment follows.

Categories of HTA

HTA was assigned to two categories by Donaldson and Sox (1992) – primary and secondary.

- Primary HTA involves direct collection of data, from or about patients, and collection and analysis of cost data. RCTs and epidemiological observational studies are included.
- Secondary technology assessment makes use of existing data; its methods include systematic literature synthesis and meta-analysis.

Cost-effectiveness and cost–benefit analyses, computer modelling, ethical, legal and social assessments are carried out at both primary and secondary research levels.

Assessment of clinical effects

The Department of Health (1992) classified HTAs of clinical effects into those that use statistical adjustment to try to control for selection biases, and those that control selection biases using randomisation. The former, 'observational' approach takes advantage of differences in clinical practice that exist between clinicians or places (cross-sectional studies) or over time (using 'historical' controls). Statistical adjustments are made for differences other than treatment differences (selection biases) between the comparison groups. In RCTs, however, the aim is to actively create unbiased groups for comparison. Participants are allocated at random to samples receiving one of two or more alternative forms of care (Department of Health, 1992).

If a technology has a large impact, observational data may be sufficient to demonstrate the effect clearly, in which case an RCT is unnecessary.

Analyses of observational data are also important in cases when sufficiently large trials are logistically and financially impracticable. The use of observational data may also raise the level of uncertainty about the effects of a technology, thus creating a climate in which RCTs are perceived to be required (Department of Health, 1992).

Drummond (1992) argued that the RCT was the cornerstone of clinical research, and that managers should be highly suspicious of evidence for the efficacy of technologies that was not generated by this approach, unless it was plausible that the impact of the new technology could be extremely large. The benefit of most new technologies is likely to be moderate, at best. Hence, the concern about using observational approaches is that their selection biases may be of the same order of magnitude, thus obscuring or grossly exaggerating any fine effects. It is widely agreed that, in most circumstances, RCTs, when rigorously executed, are by far the most satisfactory HTA method. Observational studies, which can introduce moderate biases, cannot yield reliable estimates of moderate treatment effects. In most cases, moderate effects are all that it is realistic to hope for from present treatments for common life-threatening or disabling diseases. It is chiefly to distinguish reliably between moderate but still significant beneficial effects and no effect (or mildly deleterious effects) that strict randomisation, analysis and interpretation, without bias, are used (Department of Health, 1992). Randomly allocating patients between treatments balances not just known factors but also those that are unrecognised and unmeasurable. This cannot be assumed to have been achieved in analyses of observational data, no matter how elaborate the statistical adjustments applied (Department of Health, 1992).

While RCTs are theoretically indicated in many circumstances, carrying out a rigorous evaluation by RCT may be difficult. RCTs need to be capable of reliably distinguishing between two initially plausible alternatives: either the technology confers no material benefit or it has worthwhile effects (which may be quite small) on important outcomes. This requires not just the avoidance of bias but also the minimisation of random errors; it often necessitates much larger trials than have been customary which, in turn, cost more and take longer (Department of Health, 1992). In situations in which the nature of the technology was changing during an evaluation, the reliability of the results of the RCT would come into question, raising the issue of whether the resources allocated to the RCT were justified.

Another criticism of RCTs is that they are commonly explanatory rather than pragmatic trials. They are often conducted under atypical ideal conditions, such as in specialist centres with highly committed clinical investigators who use the most advanced equipment on a highly selected and, perhaps, motivated group of patients who tend to comply with therapy, rather than among patients in everyday settings (Drummond, 1992). The trials have thus tended to evaluate efficacy (Can the technology work in an ideal setting?) rather than effectiveness (Do they work in an ordinary clinical setting?).

Black (1996) argued in favour of using a range of evaluative techniques and concluded that RCTs and observational methods should be seen as being complementary. He noted that the principal observational epidemiological methods were non-randomised trials, cohort studies (prospective and retrospective), and case–control methods. Black maintained that observational methods were needed because of the limitations of RCTs and argued that, in certain situations, RCTs may be unnecessary, inappropriate, impossible to implement, or inadequate. He considered that the problems associated with RCTs often arose from a largely uncritical transfer of a well-developed scientific method in pharmacological research to the evaluation of other health technologies and services.

In 1993, Franklin also concluded that the most important implication of the strengths and limitations of RCTs and observational studies was that neither method could by itself completely address the complexity of technology assessment; hence, the two approaches should be seen as complementary rather than competitive.

In discussing the short 'window of opportunity' that might be available for making an evaluation, Stocking (1988) suggested that opinion leaders might already have made up their minds about a technology before the results of lengthy assessments could be made available. If opinion leaders were to be influenced then 'quick and dirty' studies were needed to give at least some information about the technology and its implications.

These arguments lend support to the principle of using different methods of HTA in parallel, selected to yield a useful, combined body of knowledge on which to base decision-making. Stocking's (1988) case for an initial study at an early stage, using a quick and imprecise method of assessment to give something on which a case for further action may be built, is attractive. However, it begs the question of whether this would be a sufficiently reliable basis for subsequent decisions. Given that RCTs are expensive and that to carry them through rigorously requires limitation of the availability of a treatment and control of its format during the test period, some preliminary evaluative information is needed to reinforce the demand for this level of commitment. Stocking's model, then, suggests a pattern of clinical evaluation starting at the point when opinion leaders are becoming interested but not yet committed, with a quick, inexpensive assessment to determine what potential benefits a new technology may have to offer, followed by observational studies in combination with RCTs.

Non-clinical assessments

Dolan and Zingg (1993) pointed out that most methods of assessment focus on quantitative issues but that it is increasingly recognised that the legal, ethical and social implications of technology must be addressed as well. As with clinical assessment, non-clinical assessment poses significant methodological problems. Studying the effects of the social, ethical, legal and organisational impacts of technology requires a range of methods, both qualitative and quantitative, and a different set of skills from those needed for clinical evaluation. Much can be learned from interviewing relevant people – patients, their relatives and friends, and those providing care. Observers can also be used to assess the impact of a particular technology (Department of Health, 1992). However, achieving a satisfactory degree of rigour and acceptance by these methods is not easy.

According to Jonsson (1993), the importance of economic evaluations is increasing because such studies are relevant to both patient and physician for optimising treatment, as well as to decisions about the allocation of scarce healthcare resources. This recognition of the need to carry out economic evaluation has only recently emerged. Adams (1992) assessed the prevalence and completeness of economic analyses in RCTs published from January 1966 to June 1988 and found that, of over 50,000 published RCTs, only 121 (0.2%) included economic analyses.

Like the other types of HTA, economic evaluations are fraught with difficulties. Drummond (1992) lists four principal types of economic evaluation.

1. Cost-minimisation analysis, in which different treatments or technologies are assumed to have identical results and the aim of analysis is to establish which will achieve the outcome at least cost.

2. Cost-effectiveness analysis, in which cost of treatment is assessed in relation to each natural unit of health improvement gained by the patient (such as years of life gained or disability days avoided).
3. Cost–utility analysis, in which cost of treatment is assessed in relation to each year of life gained, adjusted for quality (i.e. quality adjusted life-year or QALY).
4. Cost–benefit analysis, in which cost of treatment is assessed in relation to benefits, quantified in terms of monetary value to the patient and to society at large.

Economic evaluation can be integrated with assessments of efficacy and effectiveness by undertaking economic assessments as part of controlled clinical trials. However, as Sculpher and colleagues (1995) noted, although the demand for economic analysis as part of healthcare research and development is growing, the best model for incorporating it into clinical research has yet to be established. They argue that an assessment of the potential for a new technology to be cost-effective needs to be undertaken at an early stage, before widespread diffusion, and that economic evaluation should be seen as a sequence of carefully planned studies. In addition, if economic evaluations are to be based on clinical trials, more extensive economic input into trial design is needed at an early stage. They describe an economic evaluation approach with four stages. If, because of the absence of external controls, new technologies are likely to diffuse rapidly, because of the incentives or pressures that face providers, then the stages of economic evaluation might need to be compressed.

Assessing different types of technology

As indicated earlier, health technologies may be developed for a variety of purposes, including drugs, devices, diagnostic techniques and surgical procedures. Feeny and colleagues (1986) stated that the strategies for definitively establishing the efficacy of a technology had become established through use in drug trials. They noted that, although it had been strongly argued that the same principles applied to other technologies, such as surgical procedures, this view was not universally accepted. Feeny and colleagues (1986) suggested guidelines both for appraising reports of the assessment of any therapeutic technology and for the appraisal and assessment of diagnostic technologies.

Banta and Luce (1993) contrasted assessments of medical imaging, surgical practice, drugs, picture archiving and communication systems. In the case of diagnostic imaging, they noted five levels of evaluation: technical evaluation, diagnostic accuracy, diagnostic impact, therapeutic impact and health impact. They concluded that the poor quality of primary data remained the main problem in assessments of diagnostic imaging.

In the case of surgery, Banta and Luce (1993) considered that the nature of surgical procedures contributed to the difficulty of testing them using RCTs, since the surgeon's skill was likely to have an important impact, and this might change over time.

With reference to picture archiving and communication systems, Banta and Luce (1993) noted that although there was little experience of technology assessment in medical informatics, this was increasing despite practical and methodological difficulties. They concluded that improvements in healthcare, as a result of computer systems, would be difficult to classify and identify, and that economic analysis, concentrating on the process of care, would continue to be the major method of assessing such systems.

Objectives

The objectives of the review were to identify principles governing the timing of HTAs, to ascertain their applicability to new and fast-changing technologies, and to assess whether the literature provided adequate guidance to inform future decision-making.

Search strategy

Systematic searches for papers on the timing of HTA were made using the databases, Medline and Embase. No papers were identified that focused primarily on the question of timing. When the question of timing was mentioned, it formed part of a wider discussion of HTA.

The products of the database searches were supplemented with articles identified through other means, which included contact with experts in the field, hand-searching the reference lists of key articles, and monitoring the contents of a small number of medical journals published during 1996. A more detailed description of the methodology appears in Appendix 1, together with the number of articles identified and principles guiding selection for analysis. Articles are listed in the References and are further classified in Appendix 4.

Studies included

Identification of new health technologies

Even before considering how to identify the time at which a particular health technology should be assessed, there is the problem of identifying if a new health technology has emerged. In a report commissioned by the UK Department of Health (1992), it was pointed out that that any arrangements for setting priorities must therefore develop an early warning system for emerging technologies.

A Dutch project with this objective was reported by Banta and Gelijns (1994). This project was set up, under the auspices of The Netherlands Government Steering Committee on Future Health Scenarios, to analyse future and emerging healthcare technology. The primary users of the project results were intended to be public healthcare policy and decision-makers. The objectives of the project were first to identify new technologies as early as possible by acting as an early warning system and, second, to assess the prospects of a number of high-priority technologies or areas of technological change.

The early identification system involved systematic scanning for indications of which technologies were to be expected, why they could be of importance and an indication of the likely time-frame for their development and introduction into medical practice. The main focus was on the identification of potential innovations while they were at the development phase. Sources of information included the published literature, news services, biomedical and bioengineering conference proceedings, as well as expert opinion on estimates of likely developments in various fields of medicine and health-care.

The conclusions drawn from the project were that achieving an early identification system that remained relevant both to operations and policy making would require a permanent structure, which would also periodically update the information collected. The most efficient way of establishing such a system was thought to be through building a network of groups, each consisting of two or three experts in various clinical and biomedical research areas, who would review the scientific and technological developments in their respective fields annually.

A Canadian body with a similar role has been created (Battista *et al,* 1995), the Canadian Coordinating Office for HTA, one of whose functions is to act as an early warning system

for emerging technologies. Information is obtained by scanning newsletters, scientific literature and bulletins, and through links with other agencies. Carrying out this task is both time- and resource-intensive. Dissemination is through publication of *Technology Briefs,* in which the aim is to address a specific technology and translate technical information for policy makers in a timely and succinct manner. This has been rated as a valuable early warning system.

A detection and control system for novel surgical techniques, proposed by the Senate of the Royal Surgical Colleges of Great Britain and Ireland, has been described by Border (1995). This was designed to address issues such as evaluation, training, continuing education and quality assurance monitoring, and was intended to operate in a series of stages.

1. Detection of new techniques through, for example, literature, communication and conference reviews.
2. Evaluation of novelty by questioning whether a new technique differed from existing practice enough to warrant assessment or training.
3. Evaluation of the procedure by clinical trials at specified centres, during which time the use of the procedure would be restricted to those centres. Trials would also develop and define training methods and requirements.
4. Introduction of the procedure, once its value had been proven, with its use restricted to those surgeons who had received appropriate training. In the event of the new procedure becoming more generally adopted, proficiency in its use would form part of basic surgical skills assessment.

In the UK, the Standing Group on Health Technology advises on national priorities for HTA. Its tasks include: identifying and prioritising technologies in need of assessment; advising when there is a particular need to control the diffusion of a technology until more information is available; identifying emerging technologies likely to have major implications for the NHS; and identifying and prioritising the need for research and development in methods used to perform HTAs (NHS Executive, 1996).

In addition, a register of new interventional procedures, SERNIP, is being established by the Academy of Medical Royal Colleges (Department of Health, 1995). This voluntary system will aim to protect patients from the inappropriate application of new interventional procedures whose safety

and efficacy has not been established. It will incorporate a method of identifying and registering such procedures and advising on how they may be evaluated in a controlled manner.

All these systems for detection of new technologies require value judgements by experts on what is 'new', and whether it is likely to give rise to health technologies that are safer/cheaper/more effective than existing treatments. In many cases, such judgements would have to be made by experts with vested interests and an 'insider' perspective; hence, there are questions as to whether the systems, as designed, would be effective or dispassionate enough to justify the resources involved. Would all emerging health technologies be detected? Would their potential be correctly assessed? Would voluntary systems work?

Timing of assessment

While there is a fair degree of consensus in the literature about the principles that should guide timing, the recommendations are couched in general terms, rather than in terms of key indicators for identifying a specific point in time at which assessment should be initiated. However, there is widespread support for the view that assessment should be on-going once the decision had been made to begin evaluation.

Assessment should be early

Bunker and colleagues (1978) reviewed the introduction and evaluation of four new operations (shunt surgery for portal hypertension, coronary artery bypass graft, jejunocolic and jejunoileal bypass and total hip replacement). They noted that there had been RCTs of three of the four operations but only some time after their introduction and after many procedures had been undertaken. Bunker and colleagues drew attention to shortcomings in the evaluations, such as a lack of systematic collection of data and reporting, and concluded that earlier trials would have speeded the process of evaluation in each case. They argued that a new treatment should be introduced in a manner that allowed prompt and reliable evaluation of its efficacy and safety. Bunker and colleagues noted that when the new treatment is an operation, its introduction is almost always uncontrolled. Thus, with increasing use, the procedures might vary over time and between sites, which might in turn lead to conflicting evidence about benefits that originally seemed clear.

The UK Department of Health (1992) held that new health technologies should be evaluated before they are allowed to diffuse into clinical practice and that, in addition to safety, effectiveness and cost issues, ethical issues should be addressed early in the development of new technologies. Gelijns and Rosenberg (1994) also considered that new technologies should be assessed as early as possible.

Banta (1986) argued that to be successful, assessments must be made early enough to affect decision-making. He pointed out that assessments were usually undertaken late in the life cycle of a product or procedure and, by that time, important decisions had already been made on the basis of personal experience or financial considerations rather than by objective assessment of the net benefits. This accorded with the criticism of Feeny and colleagues (1986) that new health technologies tended to be widely disseminated before they were rigorously evaluated to determine their clinical effectiveness. Luce and Brown (1995) also noted that results of evaluations were generally available too late, after purchasing decisions had already been made, and that opinions from speciality groups and reviews and recommendations were often published **after** the technology had become accepted practice.

An alternative viewpoint was that of McGregor (1994), who stated that new technologies needed to be fairly widely diffused before they could be assessed. He maintained that after a technology was shown to have worked, it could take years of use by many operators to develop and refine it. McGregor argued similarly that once efficacy had been proven, the technology had to be widely used before its effectiveness in general use could be evaluated. He conceded, however, that diffusion was often wider than was strictly required. This delayed approach to evaluation is open to a number of criticisms.

- Ineffective or harmful technologies would be allowed to diffuse before they were evaluated.
- Once a technology had diffused into routine practice, it would become much more difficult to conduct an adequate evaluation.
- Apparent benefits of the new treatment might dissuade physicians from taking part in RCTs and their influence, together with media coverage, might engender patient reluctance to take part in such trials.

The critical point

The concept of identifying a specific point at which an evaluation should be undertaken, rather than that it should simply be 'early', was proposed by a number of writers.

Bunker and colleagues (1978) argued that the period when the clinician developed and refined a new surgical procedure and defined diagnostic criteria for its application was akin to the beginning of a feasibility study. Independent review, when favourable, should then lead to collaborative trials. They maintained that, in the presence of the learning curve phenomenon, the timing of the shift from feasibility study to multicentre trial might be crucial. To achieve this successfully a central reviewing authority was required which was capable of sophisticated statistical and economic analysis and was empowered with the authority and resources needed to initiate and coordinate appropriate trials (Bunker *et al,* 1978).

According to Stocking (1988), the earliest group to adopt an idea are the innovators who design the new procedure or technology. Widespread adoption does not immediately occur. It is only when the opinion leaders, among the relevant group of clinicians, begin to take up the idea that the majority will follow. Stocking argues that the important period for evaluation is **the time when the opinion leaders are still considering the new technology**, since at this point there is clearly an emerging technology which has not yet diffused out of control.

Banta and Andreasen (1990) were even more specific. They considered that there was a critical point at which an evaluation should be carried out and stated that if an assessment was done too early it would be forgotten, whereas if it was done too late it would be largely worthless. They did not, however, describe the point in terms that would assist timely recognition.

On-going evaluation

In addition to proposing that there was a critical point for initiating evaluation, Banta and Andreasen (1990) argued that assessment needed to be an **iterative process** rather than a one-off study; that is, there needed to be a review of policy against a backdrop of an HTA overview. There was no established formula for the timing of technology reassessment and the optimal period between reassessments would vary, depending on the effectiveness of the technology, costs, safety and the health impact of the disease or injury to be diagnosed or prevented. Thus they again failed to offer any convenient guidelines for easy identification of critical points. They asserted that policy making often depended on understanding the implications of a new technology at several points in its diffusion. Banta and Andreasen also suggested that assessment could be thought of in terms of the

life cycle of a technology, starting with prospective assessment of a nascent technology, assessment of the impact of the technology in practice, followed by later assessments as the technology changed or was used for new indications. Later reassessment would evaluate use and determine whether replacement or abandonment should take place. This echoed the view of the (Dutch) Steering Committee on Future Health Scenarios (1987), who had argued that a complete system for assessing a new health technology should monitor change at all stages of technological development and diffusion.

A more recent report by the Institute of Medicine suggested a process for setting priorities for technology assessment, which also took into account the need to re-evaluate (Donaldson & Sox, 1992). One recommendation was that all previously assessed topics should be considered as candidates for reassessment. A change in the nature of the condition, expanded professional knowledge, a shift in clinical practice, or publication of a new, conflicting assessment might trigger consideration of a technology for reassessment. Another recommendation was that topics for reassessment should be prioritised at the same time as for initial assessment, and that the topics should be given equal weight at each assessment. This report like those preceding it described critical points but not in ways that would make them easy to pinpoint in a rigorous and objectively justifiable fashion. Identifying the critical points would require value judgements about how much of a change in circumstances is enough to warrant a new assessment. These judgements would be have to be made by experts who might themselves have vested interests and preconceptions which militated against objectivity.

ACOST, in its 1993 report on medical research and health, endorsed and expanded the views in the Institute of Medicine report. ACOST stated that if a novel development was to be promoted for general use it was essential that it should undergo scientific assessment including: evaluation of safety, efficacy and outcomes, in the short- and long-term, comparison with existing options, cost-effectiveness and indications for use. The Advisory Council recommended that the NHS should require all new medical devices or novel applications of existing devices to be developed only under controlled conditions, and to be linked to validated data collection and analysis systems in a way that would facilitate the effective dissemination of results. The development of medical advances should be seen as an iterative process in which information is collected, validated,

and disseminated, and the technology re-evaluated and refined. This would allow assessments to be made at different times in the cycle.

Franklin (1993) argued similarly that technology assessment should be an on-going process as new questions continually arise and new information is constantly being generated. Gelijns and Rosenberg (1994) also espoused on-going evaluation of continually changing technology. They noted that improvements in medicine are mostly incremental and part of a continuous process comprising numerous small-scale advances; as a consequence, the manner of use, the clinical results achieved, and the resource costs associated with technological interventions change continually.

Apart from on-going evaluation of technologies that change, there is also the question of continued monitoring of a stable procedure once it enters clinical practice. Gelijns (1990) argued that, following randomised or otherwise well-controlled efficacy and safety trials, long-term surveillance should be undertaken of the safety and efficacy of new procedures as they were used in everyday clinical practice. Such studies might involve experimental or observational methods. Bunker and colleagues (1978) also maintained that just because the efficacy of a procedure had been established by collaborative study, followed by its wider use, this should not mean that evaluation came to an end. They noted that there was a clear need for continuing observation of many procedures in order to determine their long-term results.

The validity of assessments

The authors of the papers reviewed above considered the question of timing from the perspective of the needs of decision-makers for information about health technologies at a particular point in the development/adoption process. Although there is a consensus that these technologies should be evaluated early and repetitively, there are also other questions to be addressed which relate to the rigorousness of the evaluation and its on-going relevance. These highlight the fact that the value and usefulness of the information generated is affected by the time at which the assessment is carried out.

Some of the limitations on the usefulness of the information gained from assessments were underlined by Fineberg (1985), who pointed out that the results of a technology assessment can be outpaced by fast-changing technology. He contended that, by its nature, new medical technology is in a state of flux with advances in the technical capacity of a technology to be expected, especially in the early phase after its introduction. This was also the time when the need for a technology assessment might be felt most acutely. Fineberg argued that this was only one of several difficulties faced in evaluating a fast-changing technology. Since the value of a technology relates only to the alternatives that may be available, changes in competitive technologies as well as evolution of the target technology affect its relative efficacy, safety and cost. In addition, as physicians become more accustomed to a new technology, better use is made of it even in the absence of shifts in technical performance. Added to this, the underlying disease pattern in the population to whom the technology is applied might be changing. Changes in fundamental scientific or medical knowledge that underlie perception of disease or its treatment might also alter the environment in which a technology is used. Finally, changing social values and expectations might alter the potential impact of some new technologies, for example, genetic engineering.

Fineberg's (1985) discussion underlines the extreme complexity of the environment in which health technologies are developed and applied. Lilford and Jackson (1995) focused on one of the key requirements of rigorous assessment on which rapid environmental change impacts. They argued that RCTs were only ethical under conditions of equipoise; this limits the rigour of assessment that is possible because equipoise represents a window of opportunity. Equipoise was defined as the situation in which there is no preference between the treatment options to be compared. Individual equipoise applied to individual clinicians, while collective equipoise applied to the health profession as a whole. Lilford and Jackson drew attention to the practical difficulties of rigorously adhering to the principle of equipoise, which would create problems for recruitment to clinical trials. In advance of a trial, clinicians often have rational but different preferences and consequently may not be in equipoise.

Lilford and Jackson (1995) argued that where there was a high recruitment rate it called into question the comprehensiveness of the counselling that patients had received. An ethical obligation to maximise perceived utility for individuals was likely to restrict trials that were desirable for society as a whole. This was because society as a whole needed precise answers to clinical questions and this depended on high patient recruitment rates to clinical trials. However, in circumstances where importance was attached to equipoise, clinicians, with what they believed to be a rational preference

for one type of treatment over another, would not offer trial entry to their patients. Thus it would be the minority of patients, for whom, in the face of all current information the physician was in effective equipoise, who would be considered suitable for randomisation.

One implication of the concept of equipoise is that, practically and ethically, the period during which rigorous RCTs could be carried out might be brief, since the period during which a clinician has no preference between the treatment options to be compared is likely to be short. This suggests that RCTs should be undertaken as soon as new technologies are introduced.

Chalmers (1975) argued for randomisation of the first patient, as opposed to randomisation only when the trial can be carried out without changes to the treatment. Echoing Lilford and Jackson (1995), he argued that physicians could never undertake a controlled trial if they were consciously enthusiastic about one of the procedures.

Chalmers' (1975) argument was as follows. Where surgical procedures had been confirmed by therapeutic trials, such trials had been preceded by many consecutive series of patients in which the technique might have been slowly modified. Waiting until the technique has been modified before evaluating it is unethical, however. This is because, in effect, the physician is asking certain patients to give up their right to the standard accepted therapy and to be treated instead by a procedure that has not yet been developed sufficiently to justify its comparison with that standard therapy. Chalmers (1975) maintained that it would be more ethical to randomise from the beginning and explain to patients that they had a 50% chance of receiving the more beneficial therapy. He concluded that from scientific, ethical and practical standpoints, exploration of any new therapy in sick patients should begin with randomisation into either the conventional or the new treatment regimen.

Gelijns (1990), however, pointed out that during the initial stages the practitioner's skills and expertise with a procedure are still evolving, with the result that the risks and benefits associated with the procedure may change considerably. She noted that in view of this learning curve phenomenon, the initial application of a new procedure would probably need to involve methodologically sound, non-formal experimental studies. She added that such early reporting of clinical experience might form the basis for the design of subsequent RCTs or otherwise well-controlled trials to determine a

procedure's safety and efficacy, which should be undertaken at selected institutions. This argument was also propounded by Bunker and colleagues (1978), who maintained that the benefits would include an improvement in patient outcomes after the new procedure resulting from the increasing experience of the surgeons, and more reliable statistical information through a larger number of patients receiving the procedure. The requirement for the information involved in such early reporting to be reliable lends weight to the arguments for standardised reporting procedures to be developed for both observational and experimental studies, and for technologies to be evaluated at different points on the learning curve.

Durand-Zaleski and Jolly (1990) addressed the issue of the loss of impartiality of public attitudes towards nascent technologies when high hopes had been excited. They argued that a new technology, which might offer a dramatic improvement in the health status of a group of people, should be assessed with special care. Public expectations and pressure, however, often make it extremely difficult for policy makers and healthcare providers to take the time necessary to carry out such an assessment.

A systematic review of the literature on RCTs and the problems associated with executing them satisfactorily is currently in preparation in this series by Dr Robin Prescott and others.

Discussion

The literature on the timing of HTAs is not extensive. There is a broad consensus that an HTA should be initiated early and that it should be a carefully planned, continuing process. It should take into account the wider issues of social and economic impact, as well as the clinical effects of the technology.

The authors of the papers reviewed attempted to identify the 'right' moment to begin evaluating a new health technology; because of the common assumption of safety and effectiveness by both clinicians and patients, the authors considered the difficulties of achieving an adequate degree of rigour in the evaluation process, despite inadequate evidence, either from the outset or that developed during the assessment process.

A number of the authors cited have argued that there are indeed critical points at which a first evaluation should be carried out but none have identified clear indicators that would signal the

moment. Rather, they have emphasised the complexity and multiplicity of the factors that affect the emergence of a new technology and the pattern and format of its adoption.

Identification of a critical point, before policy makers need an assessment and while there is uncertainty about the merits of the technology amongst clinicians and patients, is a pressing issue, since controlling the adoption and diffusion of a new technology becomes increasingly difficult. The other problematic area is precisely how to evaluate a particular technology at a given time. While everyone agrees, in principle, that evaluation should be undertaken, given the resource issues it will always be extremely difficult to obtain a routine commitment to evaluate without clear indicators of when and how to go about it.

The question of the indicators or triggers for first evaluation needs to be addressed. Given that it is not simply a question of following the emergence of technologies developed entirely within a medical context, a warning signal is needed to identify their emergence as sufficiently significant health technologies. The possibility of using publication trend or citation analysis to identify such critical points is reviewed in chapter 3.

Writers have experienced difficulty in trying to achieve greater specificity about when, how, and how frequently to evaluate new health technologies. One reason for this is that, with the exception of bodies addressing the problems of surgical techniques, there has been a tendency to lump together the whole range of new technologies (other than drugs) that are used to promote health, to prevent and treat disease, and to improve rehabilitation and long-term care. It is unlikely, however, that a single approach will be suited to all, although there are clearly some important common principles.

A first step towards addressing the issues would be to reduce the complexity of any one case somewhat by assigning HTAs into categories on the basis, for example, of their invasiveness and of possible characteristics such as their apparent cost advantages/disadvantages, their potential for major improvements to public health, and their ethical impact. A strongly precautionary approach would clearly be appropriate for the more invasive, ethically problematic and costly. An HTA might be assigned a score on each of several parameters and the overall total used to decide on priorities for evaluation.

A series of different tests could be applied, depending on the initial score, to decide on the next step. Provided the guidelines were sufficiently clear, this would allow decentralised care purchasers and ethics committees to decide whether they were competent to evaluate a new technology, or whether they should refer to central authority for further guidance.

The question of which methods of evaluation to apply and when to apply them would be easier to answer when the category of HTA was more precisely defined and its risk characteristics assessed. However, the general principle advocated in most of the literature reviewed here, of using a number of methods of evaluation on an iterative basis is clearly important, and its utility would be enhanced if a more unified standard of reporting in observational and experimental studies were observed. For example, the recent Consolidated Standards of Reporting Trials (CONSORT) initiative is an attempt to improve the overall standard of reporting of RCTs (Begg, 1996). Hence, a set of strong recommendations on how to carry out and record each type of evaluation would enhance the value of each. In addition, there is clearly a need to improve the effectiveness of current systems of reporting side-effects of treatments once they have become routine and the level of interest in the technologies has fallen.

The implications for policy makers are that significant new health technologies need to be identified as they emerge, possibly through some form of horizon-scanning mechanism, so that they can be prioritised and steps can be taken to assess them in an appropriate manner as early as is practicable. Otherwise, the danger remains that a combination of technology 'push', pioneer enthusiasm, and media publicity may propel the diffusion of new technologies into routine practice before they have been adequately evaluated. The implications for healthcare managers and clinicians are that, where the introduction of new health technologies is not covered by central regulation, national initiatives such as SERNIP should be supported. In areas of medicine where no such national initiatives exist, the establishment of local guidelines would help to ensure that new health technologies were considered in a structured manner before any decision to adopt them. Healthcare managers and clinicians also need to be aware of the results of assessments that may have already been carried out on new health technologies whose introduction they are considering.

Chapter 3

Bibliometric study

Summary

Measuring publication trends through correlating dates of publication, numbers, authors and content of articles, identified by searching literature databases, is a speedy method of scanning the nature, and level of a reported activity. The accuracy of this method is limited by a number of factors, including the choice of search terms, the quality and coverage of individual databases and the conventions of reporting within a literature. Databases of popular press reports are likely to be particularly prone to yield misleading results owing to the unsystematic nature of journalistic reporting and the much less precise use of language compared with technical literature.

With those limitations in mind, publication trends over time in relation to each of the six medical applications were examined to see whether this approach might yield at least a crude indication of points at which policy makers should be alerted to the need to initiate evaluation. The starting assumption was that publication trends might in some way be related to the rate of development and diffusion of those applications, and that it might be possible, by plotting trends, to identify the characteristics of points at which specific evaluative activities should be triggered. Rate of citation was also investigated for the same reasons. Although this approach yielded interesting and suggestive results, it did not lead to the identification of specific 'critical points' which might have acted as indicators to trigger first evaluation.

Background

It would be extremely useful to have indicators, even if crude, of the best time to initiate HTAs in order to provide for rigorous evaluation of new health technologies before their widespread diffusion. The hypothesis was proposed that publication and citation trends might show characteristic patterns that could be used to identify such 'critical points'. Identification of key points within such a pattern might help to control adoption and might justify the resources required to carry out an evaluation.

Objectives

The objectives were:

- to display graphically publication and citation trends for the six applications reviewed
- to relate the patterns generated to events in the development and diffusion of the health technologies
- to use the information generated to confirm or disprove the view that there might be characteristic inflection points in the curves generated, which could in turn be used as indicators for stages of development and diffusion.

Search strategy

The search strategy is described in Appendix 1.

Methods

Bibliometric approaches are becoming more powerful as increasing numbers of large databases of publications become available on CD-ROM. They have now been used in a variety of ways as indicators of research and development activity.

Rappa and Debackere (1992a;b) argued that electronic bibliographic databases could be used as a source of data for monitoring technological progress. Rappa and Garud (1992) used the scientific literature in this way to model the contribution spans of scientists in the field of cochlear implants. Rappa and colleagues (1992) also employed the scientific and patent literature as a source of data to analyse the relationship between author/inventor contribution spans and the rate of technological progress in two chemical fields. Santarelli and Piergiovanni (1996) carried out an analysis of literature-based innovation output indicators, as did Coombs and colleagues (1996).

Results

The numbers of publications per year identified from Medline for each of the six medical applications under consideration are shown in *Figure 1.*
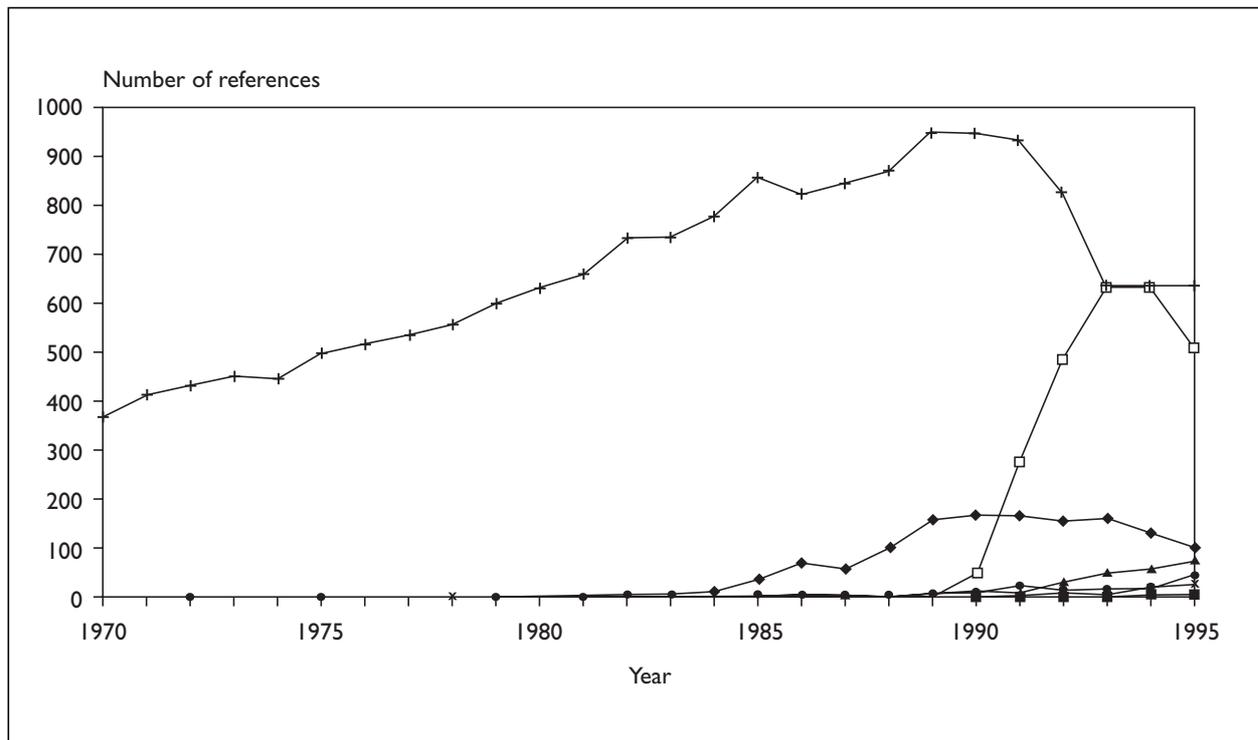
FIGURE 1 *Medline references by year for the six medical applications and coronary disease (x 0.25) (–◆–, CVS; –x–, genetic screening for breast cancer; –+–, coronary disease; –□–, laparoscopic cholecystectomy; –■–, teledermatology; –▲–, gene therapy for cystic fibrosis; –●–, teleradiology)*

The number of publications on coronary disease has been used as a control against which the publication trends of the six medical applications may be compared, on the assumption that research in coronary disease, a major long-term area of research, could be used as an indicator of the general level of medical research output. The terms used in the searches are listed in Appendix 1.

As the number of references retrieved for coronary disease was large enough to 'flatten' the trend lines of the other applications on the graph, the total number of references for coronary disease was divided by four. This had the effect of scaling down the curve while still showing the trend line and allowing the publication trends for the remaining six medical applications to be clearly displayed.

The trends for the six applications considered in this report are quite different from the trends for coronary disease and can be assumed to show research trends specific to the applications themselves. While annual numbers of publications for the relatively established applications of laparoscopic cholecystectomy and CVS appear to have peaked and are now declining, those for the other four applications are increasing. This tends to confirm the starting assumption that the first two applications have been relatively well-evaluated

and are no longer such active subjects of research and evaluation, while demonstrating that the other applications are still at a relatively early stage in the adoption and diffusion process. Teleradiology, for which references have appeared in small numbers over a number of years, illustrates the wide variations that can occur in the time-span of the adoption and diffusion process.

The number of publications per year identified from Medline for laparoscopic cholecystectomy and the period over which some of the main RCTs of the procedure took place are shown in *Figure 2*. Numbers of publications are represented by the bars on the chart, and each of the RCTs is represented by a line (which relates to the *x* but not the *y* axis).

A single reference appeared in 1989, after which annual numbers increased steadily, peaking over the period 1993/94 at over 600 references per year; the beginning of a decline in publication numbers appeared to be indicated in 1995. Thus, over a relatively short period, a significant amount of publishing activity was generated. The apparent decline in the publication rate suggests that this procedure has now become relatively routinised.

Three of the RCTs shown compared laparoscopic with open cholecystectomy and five compared
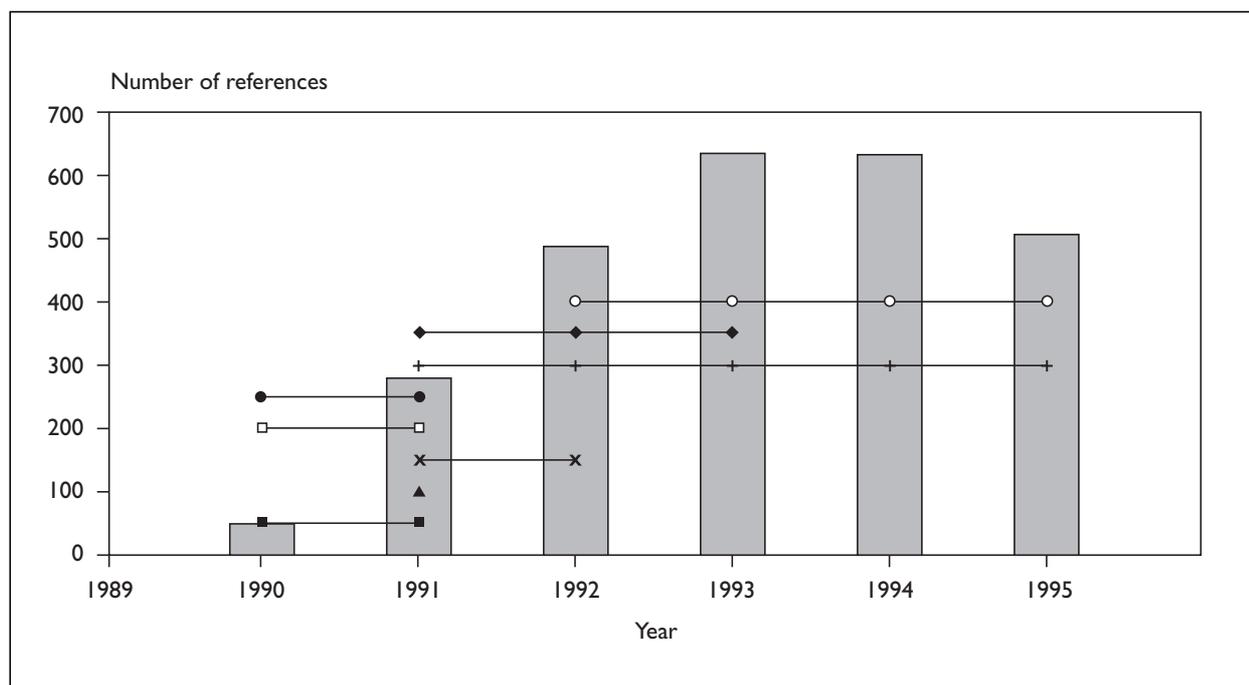
FIGURE 2 *Laparoscopic cholecystectomy (LC) – Medline references, showing duration of selected RCTs versus mini- (MC) or open (OC) cholecystectomy (▦, laparoscopic cholecystectomy; ⨯ Jan and Chen, LC–OC; + McGinn, LC–MC; ■ Trondsen, LC–OC; –□– Barkun, LC–OC; ◆ McMahon, LC–MC; ▲ Berggren, LC–OC; ● Kunz, LC–MC; –○– Majeed, LC–MC)*

laparoscopic with mini-cholecystectomy. The earliest trials shown here began in 1990 (Trondsen, laparoscopic versus open; Barkun, laparoscopic versus mini; Kunz, laparoscopic versus mini), while the latest began in 1992 and was still on-going in 1995 (Majeed, laparoscopic versus mini); all the trials began while annual publication rates were still increasing. There has been no large-scale RCT comparing laparoscopic with open chole-cystectomy and, by 1991, it was being suggested that the obvious benefits of the laparoscopic procedure had placed ethical constraints on the undertaking of such a trial (Neugebauer *et al,* 1991). Nevertheless, as *Figure 2* shows, smaller RCTs comparing laparoscopic with open chole-cystectomy (Berggren *et al,* 1994) were begun in 1991. The chart shows that the RCTs were not initiated until 3 years after the first Medline reference to laparoscopic cholecystectomy and, by this time, there were reports of ethical prob-lems about undertaking full RCTs. This suggests that a 'window of opportunity', which may be connected to the pattern of publication trends, opens, during which it is possible to initiate an RCT. After a time, this window may close and it then becomes much more difficult to undertake such evaluations. In the case of laparoscopic cholecystectomy, that window was probably some time before the first trials, since at the time they started there were already difficulties with randomisation.

The publication trends for laparoscopic cholecystectomy in medical literature (Medline), popular literature (Promt) and technical literature (Compendex) are shown in *Figure 3.* Searches were made on these databases and the results plotted, in an attempt to identify the emergence of any significant patterns in publication trends. As might be expected, coverage of this technique in the popular literature began around the same time as reports began to appear in the clinical literature. From 1990, however, while reports in the clinical literature increased rapidly until 1993, popular coverage increased only slightly until 1992, before declining from even that modest level. Thus it appears unlikely that media reports played a major role in influencing the timing of trials of this appli-cation. This surgical procedure was barely men-tioned in the technical database, which might be suggestive of the fact that the technology itself was not novel but had been introduced into other applications in the health sector through a process of technology transfer. MIS was already in use by gynaecologists and direct transfer occurred. Again, as might be expected, by far the largest number of publications on laparoscopic cholecystectomy occurred in the clinical literature.

The annual numbers of references to CVS identified through a search of the Medline data-base are shown in *Figure 4.* Also shown are the annual numbers of conferences at which CVS
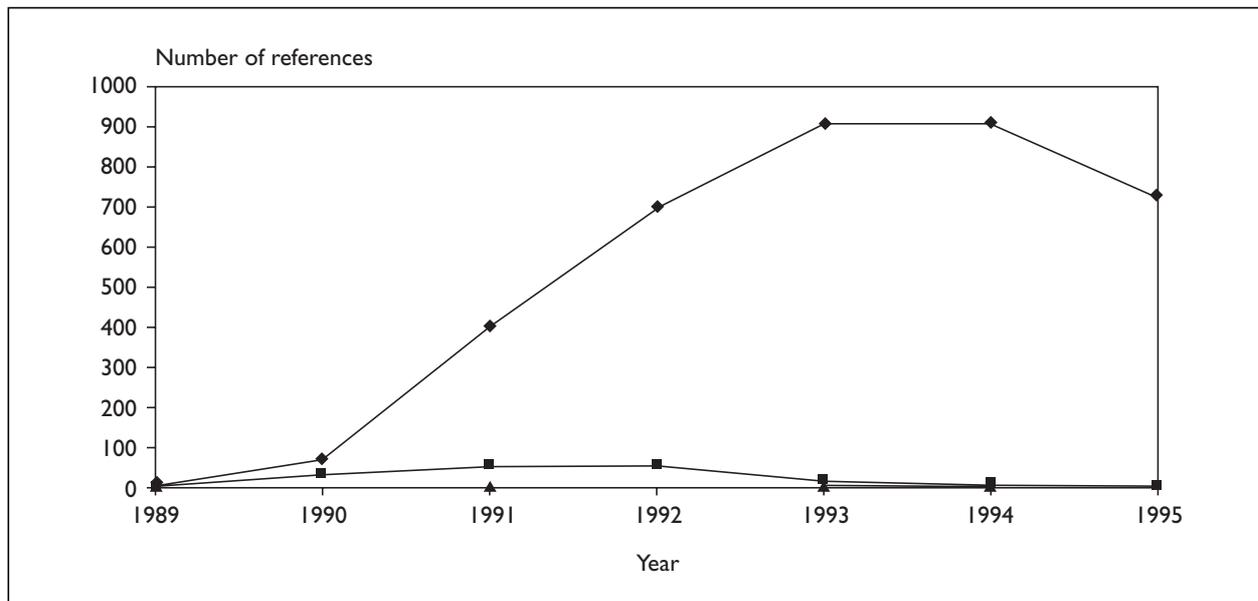
FIGURE 3 *Laparoscopic cholecystectomy – publication trends (–♦–, Medline; –■–, Promt; –▲–, Compendex)*

was featured, identified from a search of the Index of Scientific and Technical Proceedings and Medline. In addition, the major clinical trials which were undertaken comparing first-trimester CVS with second trimester amniocentesis are included (represented by lines parallel to the *x* axis) – the Canadian collaborative trial 1984–88 (Lippman *et al,* 1988), the Danish trial 1985–90 (Smidt-Jensen *et al,* 1992), the MRC European trial 1985–89 (MRC Working Party on the Evaluation of Chorion Villus Sampling, 1991) and a large US trial 1985–86 where randomisation was attempted but had to be abandoned (Rhoads *et al,* 1989).

References to CVS first appeared in the literature in 1983, gradually increasing in numbers and then reaching a plateau for the years 1989–93; since then they have been declining steadily. This suggests that CVS has now been available long enough to have become either relatively well evaluated and routinised or alternatively evaluated but not adopted widely. It should be noted when comparing publication trends for laparoscopic cholecystectomy and CVS on Medline, the highest annual number of references to laparoscopic cholecystectomy was 633 in 1994, whereas the corresponding figure for CVS was far fewer – 167 in 1990 *(Figure 1).*

Most publications on CVS followed the start of the major trials, in contrast to laparoscopic cholecystectomy where publications and trials ran in parallel to a large extent. The Canadian trial began in 1984, the year after the first references to CVS appeared, with the other major trials starting in 1985. One remarkable aspect of the Canadian trial was that CVS was only available to women who participated

in the study. One of the factors that may have contributed to this achievement was that the trial was begun so soon after the first CVS references began to appear in the clinical literature, when it could be argued that there would still have been much uncertainty over issues such as safety, efficacy and diagnostic accuracy. Conversely, the later starting date of the Danish trial may have contributed to the difficulty of recruiting patients for randomisation, with CVS also being available at centres in Denmark other than those participating in the trial.

Most of the data on conferences featuring CVS were obtained from the Index of Scientific and Technical Proceedings. Although it is a representative rather than exhaustive list, it nevertheless provides a useful indication of the numbers of conferences taking place over this period, at which papers on CVS would have been delivered, the technique discussed and information exchanged. *Figure 4* shows one conference in 1983, four in 1984, six in 1985, with the numbers from 1986 to 1992 varying from three to five conferences annually.

It appears from *Figure 4* that, as appeared to be the case with laparoscopic cholecystectomy, there was a rather narrow window of opportunity when it was possible to initiate an RCT of a new procedure. This window would appear to have been relatively brief (arguably less than 2 years in duration), beginning when papers on the new procedure first started to appear. By the end of this period, coverage of the topic in the clinical literature and the wider media might have come to influence both clinicians' and patients' views, to have reduced perceived uncertainty about the new procedure,
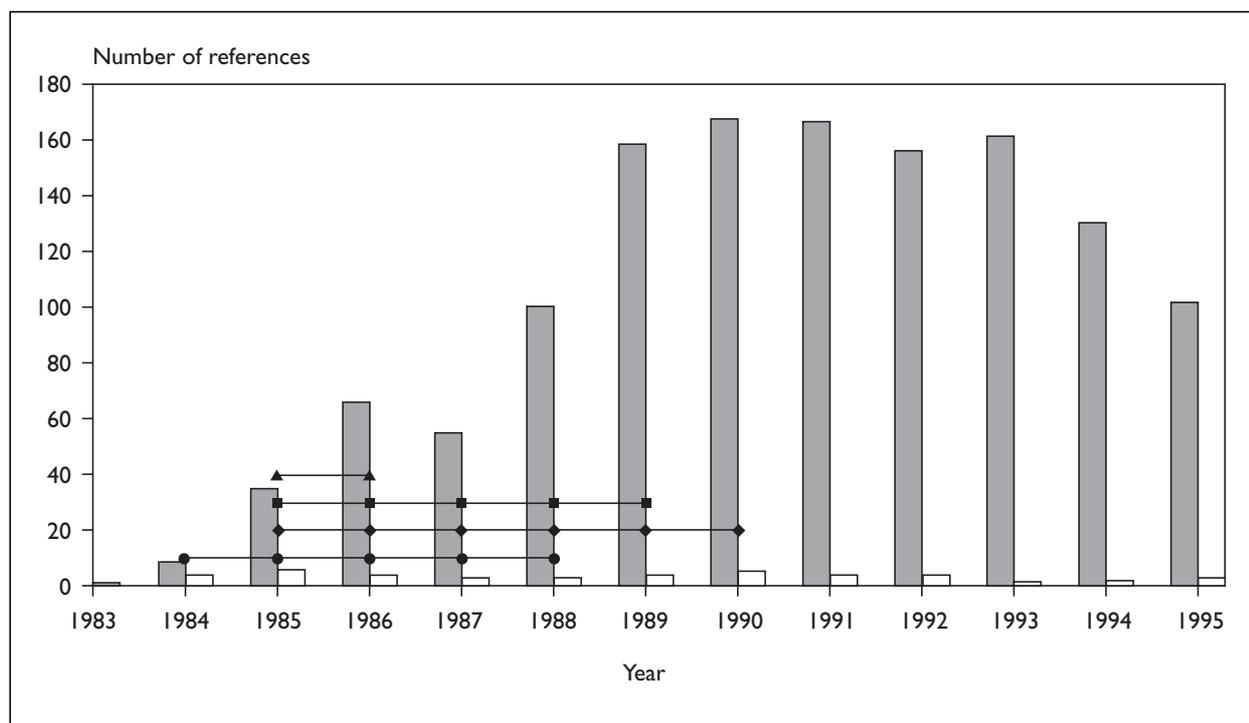
FIGURE 4 *CVS – publication trends, conferences, major trials (▭, CVS references; ▢, CVS conferences; –●–, Lippman et al, 1988; –◆–, Smidt-Jensen et al, 1992; –■–, MRC Working Party, 1991; –▲–, Rhoades et al, 1989)*

and to have made it increasingly difficult to initiate an RCT.

The publication trends for CVS in the medical, popular and technical literature (as represented by Medline, Promt, and Compendex, respectively) are shown in *Figure 5.* As with laparoscopic cholecystectomy, the largest number of references occur in the medical literature, with little mention of it in the technical literature. Although the first reports of CVS appeared in the medical and popular literature in 1983, it did not feature in the popular literature again until 1986, which saw the first peak for reports of the procedure in the medical literature. This might suggest that the increasing coverage in the medical literature led to renewed interest in the popular literature, rather than vice versa. The absolute number of references in the popular literature was very small but the results of searching the Promt database may be less reliable than those obtained from Medline.

Laparoscopic cholecystectomy and CVS have been described as the more stable and well-evaluated applications. The genetic and telemedicine technologies are representative of newer, faster-changing applications. The publication trends for gene therapy for cystic fibrosis in the medical, popular and technical literature (Medline, Promt and Compendex, respectively) are shown in *Figure 6.* The first references to gene therapy for cystic fibrosis

appeared in the medical literature in 1990, increasing steadily from 1992 onwards. Despite this steadily upward trend, however, the overall number of references was relatively small, at less than 80 for 1995.

Unlike laparoscopic cholecystectomy and CVS, the publication trend for popular reports relating to gene therapy for cystic fibrosis based on Promt was not dwarfed by that of the medical literature. Reports in the popular press seem to have appeared in 1984–85, while the first report retrieved by the Medline search appeared in 1987. In addition, the numbers of publications in the popular literature were greater than in the medical literature for 1991–94, especially in 1992. It was not until 1995 that the numbers of references in the medical literature exceeded those in the popular literature, although there were more references in the medical literature overall. As with the previous applications, references from the technical literature (as represented by the Engineering Index), were negligible. The absolute number of references in the popular literature was much higher than for laparoscopic cholecystectomy or CVS.

Why is gene therapy so newsworthy? Perhaps partly because genetic technology is perceived as pushing back the frontiers of science, and because it raises important ethical and social issues for the population as a whole? The influence of media coverage on clinicians and patients might have led
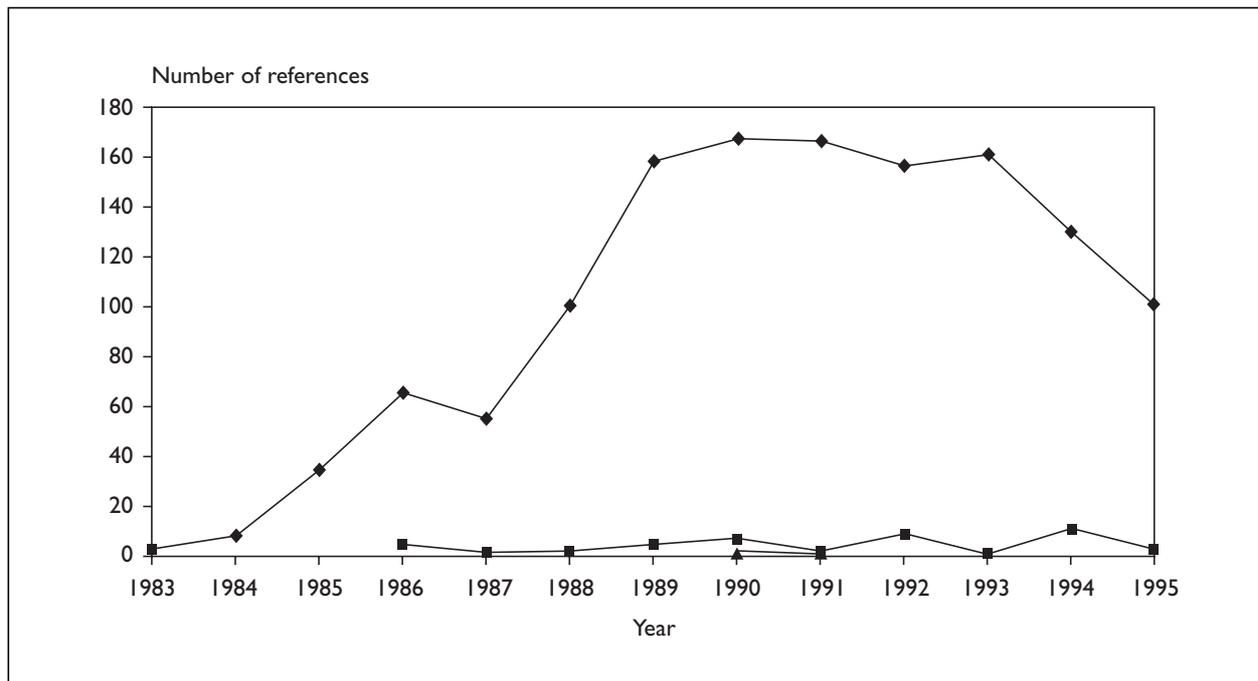
FIGURE 5 *CVS – publication trends (◆, Medline; ■, Promt; ▲, Compendex)*
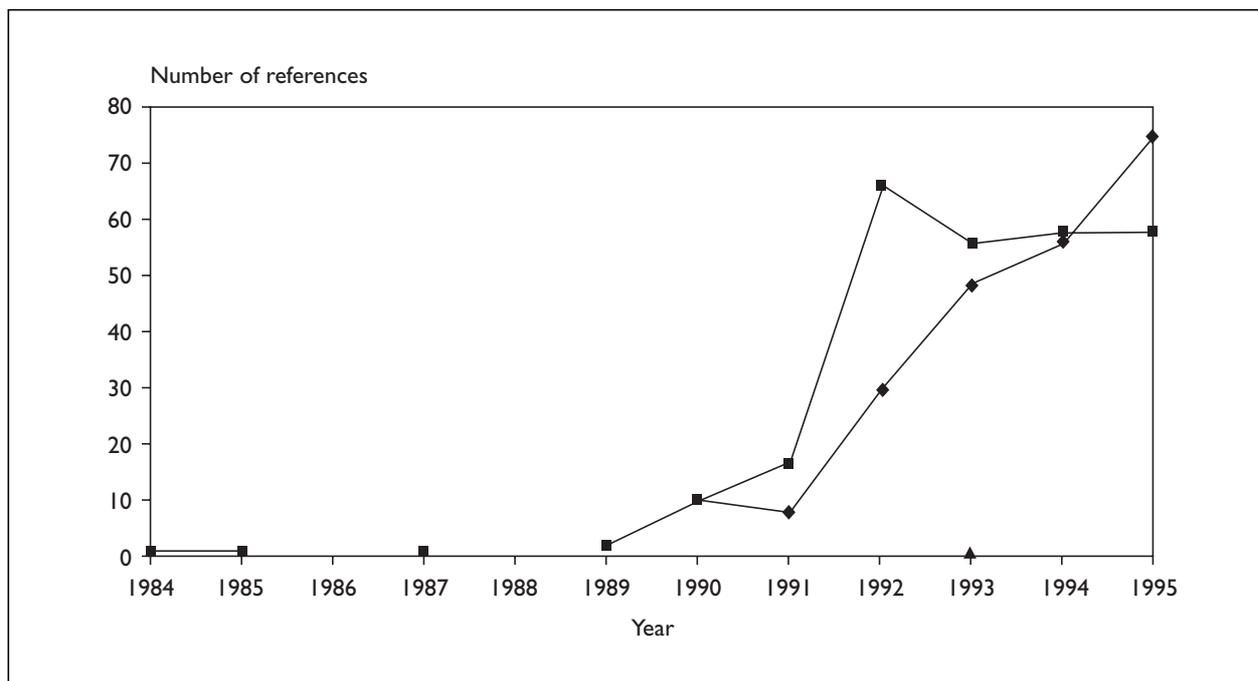
FIGURE 6 *Gene therapy for cystic fibrosis – publication trends (◆, Medline; ■, Promt; ▲, Compendex)*

to difficulties in initiating RCTs but for the fact that its development was tightly centrally controlled, in contrast with the development of laparoscopic cholecystectomy. This is discussed further in the chapters reviewing these applications.

The publication trends for genetic screening for breast cancer in the medical (Medline) and popular (Promt) literature are shown in *Figure 7.*

(No references were retrieved by the Compendex search.) The Medline search strategy retrieved one early reference to the subject in 1978 (Purtillo *et al,* 1978), which discussed the basic genetic mechanisms responsible for tumour formation. Other than this, references first began to appear in 1990 and rose steadily, apart from a dip in 1993. Total annual numbers of references in the medical literature were still low relative to gene therapy,
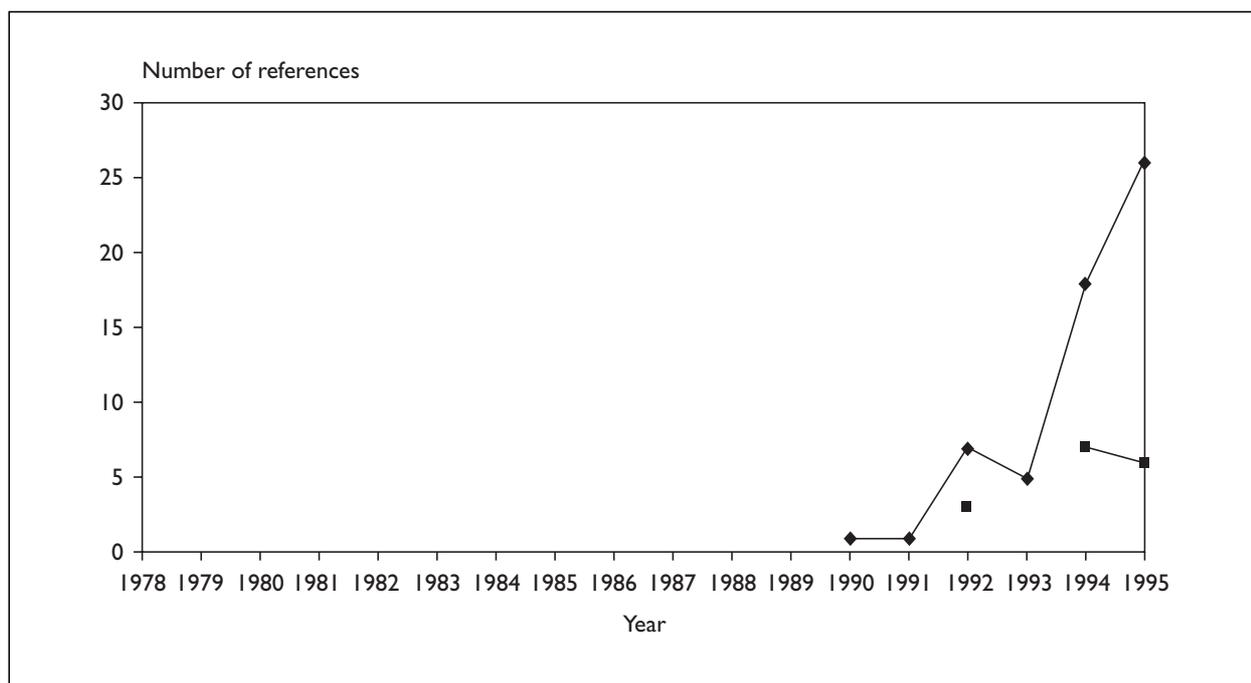
FIGURE 7 *Genetic screening for breast cancer – publication trends (◆, Medline; ■, Promt)*
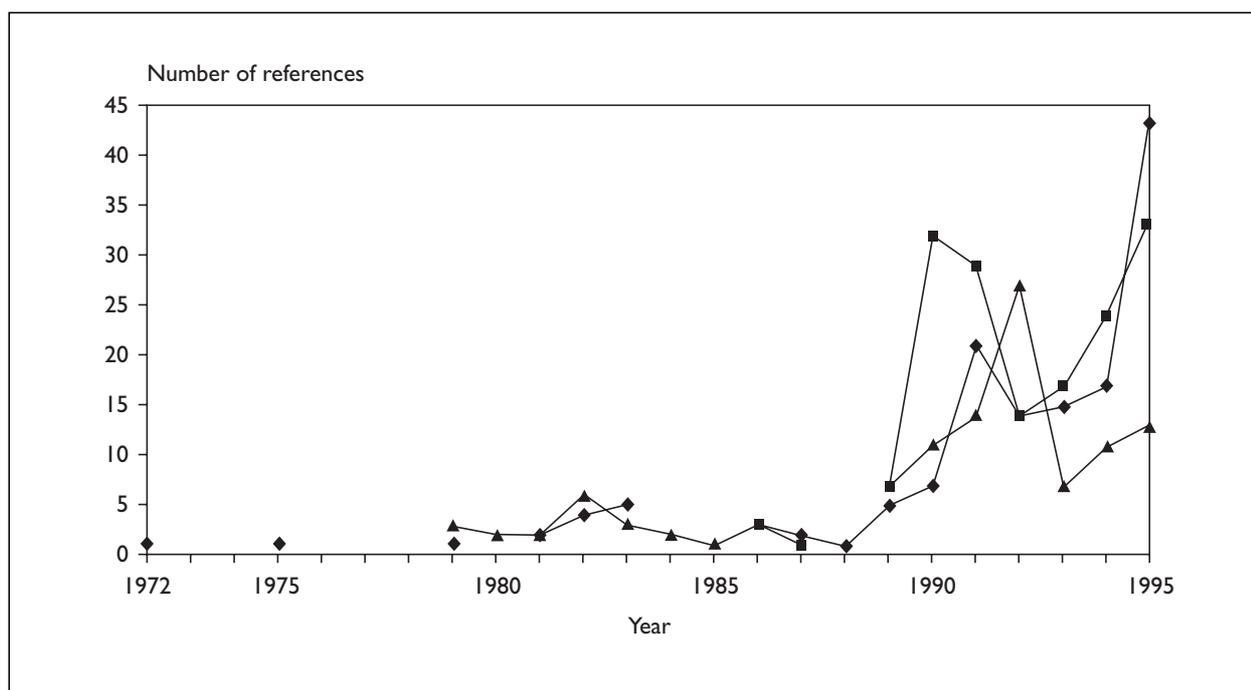
FIGURE 8 *Teleradiology – publication trends (◆, Medline; ■, Promt; ▲, Compendex)*

however, with the figure for 1995 standing at 26 (as at October 1996). Popular reports first appeared in Promt in 1992; there was a gap in 1993, then more reports in 1994, although the numbers were also very small. (This may be at least partly because the search on Promt was for material specifically relating to genetic screening for breast cancer, rather than to genetic screening in general.)

The publication patterns of teleradiology *(Figure 8)* contrast with those of the earlier applications. In the medical literature, the first reports appeared as early as 1972 but numbers of references remained relatively low up to about 1990, with a generally increasing trend from then onwards, culminating in a major increase for 1995 (although at 43 references the number

was still relatively low compared with the stable applications). This variable pattern reflected the cyclical development of telemedicine in general, in that roughly once every decade advances in the generic technology would generate new activity, which subsided when funding ran out and projects could not be sustained. The upward trend from 1990 reflects the current resurgence of interest in telemedicine as a result of further technological advances combined with greatly reduced costs in infrastructure developments and, hence, the generally higher feasibility for cost-effective use in healthcare applications.

The popular literature on teleradiology, as with gene therapy, displays a higher profile relative to the medical literature than was the case for either laparoscopic cholecystectomy or CVS. Here the newsworthiness of the topic may relate to the general interest in technological advances which have the potential to change the pattern of health-care delivery radically. It may also reflect interest in the breadth of applications of telecommunications technology rather than specifically medical applications, or may reflect the coverage of the Promt database.

Unlike any of the other applications, the number of publications in the technical literature for tele-radiology is relatively high. This may be because advanced telecommunications equipment is central to the operation of teleradiology and the technology was developed outside of the health sector

and introduced through a process of technology transfer. Alternatively, it may again reflect a bias in the coverage of the database, Compendex.

All three types of literature display a somewhat similar pattern of publication from about 1989 onwards, with initial increases being followed by dips, and then further increases.

The publication trends for teledermatology (shown in *Figure 9*) show the fewest references of all the applications, with the Medline search retrieving one reference in 1992, two each in 1993 and 1994, and four in 1995. No references were retrieved from the popular or technical literature searches. This suggests that teledermatology is a much more recent development than, for example, teleradiology, with a much lower level of activity. Alternatively, the current level of teledermatology research and diffusion may not be adequately reflected in the clinical literature covered by Medline or Embase. The Telemedicine Information Exchange on the World Wide Web does carry a significant number of references to teledermatology. This could be because its main use is in the primary care sector where there is less focus on publication in mainstream journals.

Citation analysis

In addition to the analysis of publication trends, the possibility of using the citation rates for key publications was also investigated on the basis that this too might be related to the rate of diffusion of
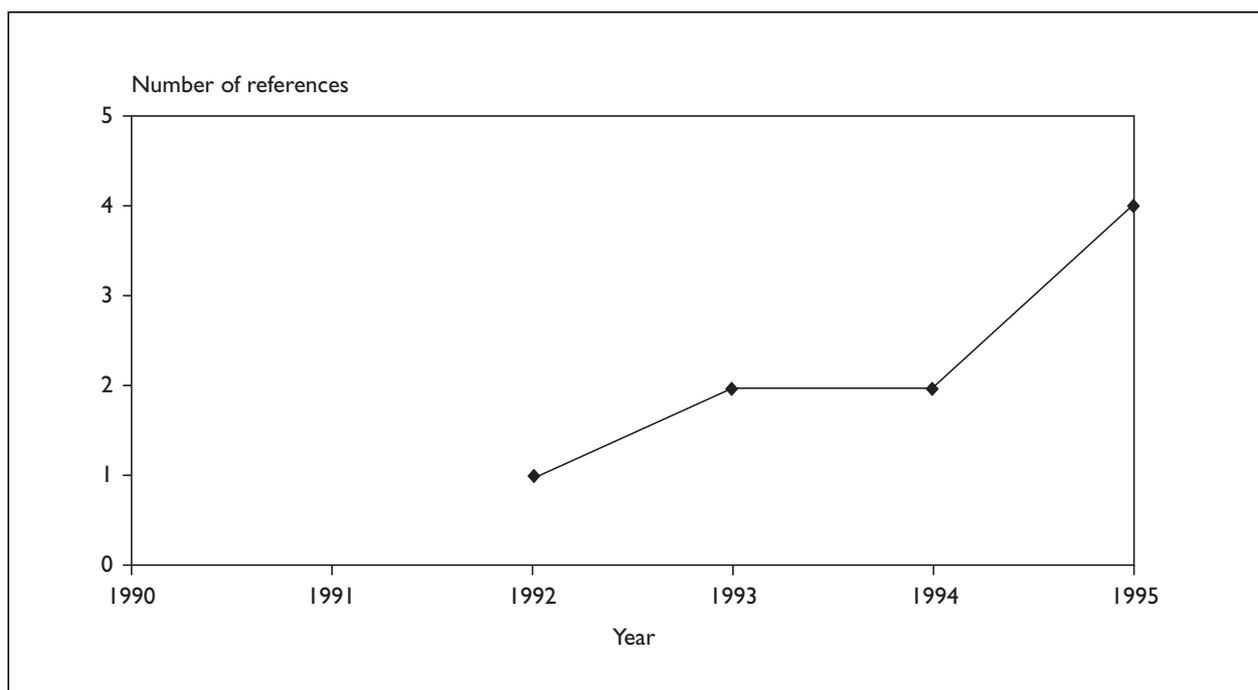
FIGURE 9 *Teledermatology – publication trends (—◆—, Medline)*

Number of times reference cited
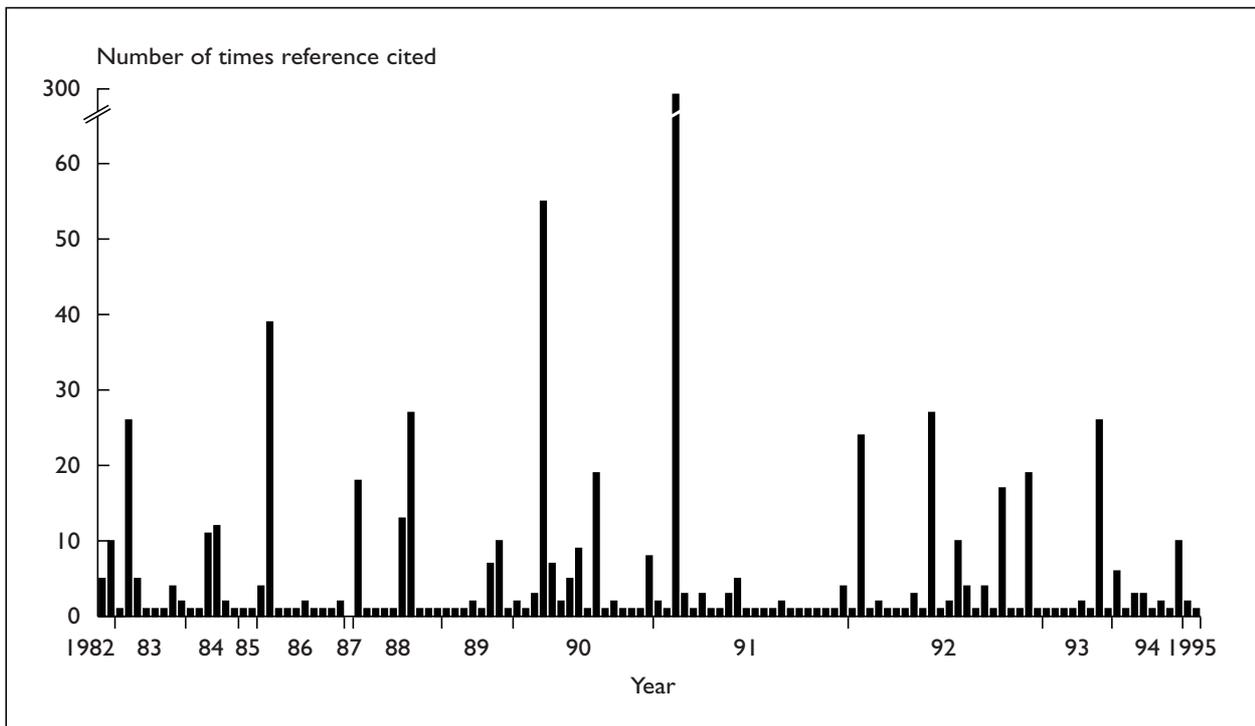
FIGURE 10 *Science Citation Index – papers by Cuschieri (and others)*

medical applications. It might also be possible to identify the key points at which evaluation might be instigated.

Information derived from the Science Citation Index is presented in *Figure 10* on the citation of papers published by Cuschieri (including papers of which he was co-author), a leading proponent of laparoscopic cholecystectomy and responsible for the first such operation in the UK on humans in 1989. The *x* axis shows each paper published with Cuschieri as an author for the period 1982–95 that has been cited by other authors. The *y* axis shows the number of times each paper has been cited in total. The papers are not necessarily restricted to the subject of laparoscopic cholecystectomy.

Assuming that rates of citation in the literature for a given application may be in some way related to the diffusion of that application, it appears that the paper published in 1991, cited nearly 300 times, might be a key paper in the literature on laparoscopic cholecystectomy. This journal article by Cuschieri and colleagues reported a retrospective survey of a number of evaluations made at European centres where laparoscopic cholecystectomy had been performed and was a review rather than a report of primary research. Thus, it seems inappropriate to interpret the large number of citations of this paper as providing an indicator for initiating an evaluation of the procedure. Indeed,

even in 1991, it was being argued that it was no longer possible to undertake an RCT comparing laparoscopic and open cholecystectomy because of ethical constraints (Neugebauer *et al,* 1991).

Discussion

These bibliometric approaches yielded interesting and suggestive results, although no specific 'critical points' were identified that might have been used as indicators to trigger first evaluation of a health technology. Certain applications, such as tele-dermatology, have as yet a very limited literature, so that the information conveyed by their publication trends may not be particularly meaningful.

In the case of laparoscopic cholecystectomy and CVS it appears that the window of opportunity for starting an RCT occurred during the time of the initial rapid rise in citation numbers. From the limited evidence presented here this period seemed to extend from the time that the first paper appeared in the clinical literature until 2 or 3 years later. In the case of laparoscopic cholecystectomy, the papers reviewed in chapter 4 indicate that the initiation of trials was delayed beyond that point, and that full randomisation was no longer possible by the time they started. Trials of CVS were started promptly and completed before the adopting group had widened. This is evidenced by the much later growth of publication about CVS relative to

the initiation of trials. However, it would be necessary to plot the trends of a large number of stable applications in order to confirm whether this 'critical point' was generally applicable. The rather different adoption histories of the now widely used laparoscopic cholecystectomy and CVS (which, after positive results in trials, was shown to have rare, serious side-effects that have limited its use to women at high risk of carrying a foetus with a chromosomal abnormality) must also be borne in mind. They are discussed further in the reviews of these applications.

A number of limitations to the bibliometric approach are also acknowledged. The data obtained were of a representative rather than an exhaustive nature and dependent on the search strategies used and the holdings and scope of the selected databases. The case of teledermatology, where more references were retrieved from a specialist database rather than from Medline, indicated that some applications might be covered in a different subset of the technical literature than those covered by the databases used. As the analysis was based on volume of publications, some documents which were irrelevant might have been retrieved by the search strategy. Also, publication occurs at varying lengths of time after the **start** of a study – and it is the starting point which is of greatest importance for timing. Finally, volume of publications and citation patterns by themselves might not provide an adequate reflection of the numerous factors that contribute to the diffusion of a new medical application.

Although the basis of selection used in the current study did not yield conclusive results, it would be worth extending this approach to identify better selection criteria and, perhaps, a wider range of literature to search. By this means it might be possible to create a useful if fairly crude tool to assist in prospective identification of key points.

Chapter 4

Systematic review of diffusion and evaluation of laparoscopic cholecystectomy, focusing on factors influencing timing

Summary

The purpose of this systematic review was to investigate when and how evaluations of laparoscopic cholecystectomy have been undertaken and to clarify the factors influencing the timing of those evaluations.

Laparoscopic cholecystectomy is a form of minimal access surgery (MAS) which has diffused rapidly into routine clinical practice. Factors influencing the rate of diffusion included technological 'push' from manufacturers of the instrumentation, and the 'pull' of demand, both from health professionals and patients, and in the early stages there was no requirement for surgeons to undergo specialised training before adopting the procedure. The technique has now been widely adopted, to the extent that the rate of publication of evaluations of the procedure has peaked and is now declining. The diffusion of laparoscopic cholecystectomy was not preceded by adequate evaluation of the technique. Many of the assessments in the literature are uncontrolled descriptive studies. There has been no large-scale RCT of laparoscopic versus open cholecystectomy; a few small RCTs have been undertaken and there have been a similar number of RCTs of laparoscopic versus mini-cholecystectomy. The observational studies and RCTs of laparoscopic versus open cholecystectomy generally reported a longer operation time, a shorter hospital stay, less postoperative pain, a faster return to normal activity and a much smaller scar. The most serious complication was an increased risk of injury to the bile duct, with a range of 0–4% in the RCTs and 0.6–1.8% in cohort studies and case series (Downs *et al,* 1996). Early RCTs of laparoscopic versus mini-cholecystectomy found in favour of laparoscopic cholecystectomy, although later RCTs concluded that it offered no clear advantage.

Few RCTs included any economic analysis. Those economic studies which have been undertaken suggest that laparoscopic is less costly than open but more costly than mini-cholecystectomy. The situation is complicated by the fact that costs seem to vary depending on the site and do not take into account the changing pattern of demand since the introduction of laparoscopic cholecystectomy.

The timing of evaluation of a new health technology is critical. There is a window of opportunity when it is possible to conduct a randomised trial, after which it becomes much more difficult owing to diffusion of the technology. In retrospect it is clear that systematic evaluation of laparoscopic cholecystectomy should have started from the earliest stages of its introduction but this did not happen. One reason for this was that there was no consensus to withhold new techniques until they were properly evaluated, neither was there centralised control to prevent diffusion of unevaluated techniques. Favourable clinical audit during the 'learning curve' period resulted in ethical objections to conducting an RCT, while the exact role of the RCT in the evaluation process was debated. In addition, media reports fuelled initial demands for the new treatment. Ironically, these also eventually highlighted complications resulting from surgeons adopting it without adequate training.

Background

Traditional surgical procedures involving internal organs consist of the three main stages of cutting the patient open, removing or repairing an organ or tissue, and closing the patient up again (ACOST, 1993). MAS reduces the impact of the first and last of these stages as far as possible, either by gaining access to the body through natural orifices or by operating through very small incisions (Border, 1995). Laparoscopic surgery is a form of MAS involving operations carried out through incisions made in the abdominal wall.

The basic technology required for such operations is described by Border (1995); it comprises imaging, surgical and other medical equipment, as well as access to back-up facilities. The vast majority of MAS techniques rely on video images, which provide high-quality, magnified colour pictures

and are displayed on two or more high-resolution TV monitors to ensure that all members of the operating team can see what is happening. The development of fibre-optic light sources, miniature video cameras and specially designed surgical tools has allowed increasingly complex surgery to be performed, with the surgeon guided by high-resolution, magnified, video images. This contrasts with the use of conventional surgical instruments guided by the surgeon with a direct view of the instruments and the body parts.

Cholecystectomy (removal of the gallbladder) is a long-accepted method of treating patients with symptomatic gallstones. Open cholecystectomy was first performed by Langenbuch in 1882 (Macintyre & Wilson, 1993). Downs and colleagues (1996) describe the three current approaches to cholecystectomy. In open or traditional cholecystectomy a 10–15 cm incision is made, in mini-cholecystectomy an incision of 5–7 cm is made, while in laparoscopic cholecystectomy three or four incisions, varying from 0.5–1 cm are made to provide access for the laparoscopic and surgical equipment and an opening through which the gallbladder is removed.

Traditional cholecystectomy had established the clinical effectiveness of removal of the gallbladder for symptomatic gallstones. What required evaluation, therefore, was the safety, efficacy, effectiveness and cost-effectiveness of the new technique compared with those of the traditional technique. Would surgeons be able to perform the same internal surgery successfully using different, remote-controlled tools, and utilising remote rather than direct visual information, as compared with the established open technique? How would the new procedure affect operation time, postoperative pain, length of hospital stay and recovery time? How acceptable would the procedure be to patients? What would be the costs, and the implications for training and for the wider hospital environment? These were some of the questions that should be answered through evaluation of the new procedure. This review aims to clarify the factors that influenced when questions such as these were addressed, and the methods used to address them.

Objectives

The objectives of the review were to identify when and how evaluations of laparoscopic cholecystectomy had been undertaken and to clarify the factors which influenced the timing of those evaluations.

Search strategy

The databases, Medline and Embase, were searched for papers on laparoscopic cholecystectomy, using this as the primary search term; the search period was from 1989, when the first laparoscopic cholecystectomy was performed, until September 1996. The database searches were supplemented by other means, including contact with experts in the field, hand-searching reference lists of key articles, and monitoring the contents of a small number of medical journals in 1996. (See Appendix 1 for more details of the search strategy.)

Studies included

Development and diffusion

Many medical applications involving the use of technological equipment only become feasible when the technology has been sufficiently developed to allow its application within the field of medicine. Such development often occurs outside the health sector and the new medical technology is introduced through a process of technology transfer. Szczepura and Kankaanpaa (1996) noted that laparoscopes were available in the 1960s but that the imaging systems and instrumentation were not of a sufficient quality to allow their use in therapeutic interventions. The refinement of high-resolution video cameras and the development of appropriate instruments eventually made their adoption for medical applications possible.

The French surgeon Mouret is generally credited with having performed the first laparoscopic cholecystectomy on humans in 1987 (Cuschieri *et al,* 1991). By the following year several groups were independently developing the technique: Dubois and colleagues in Paris; Perissat and colleagues in Bordeaux; Reddick and Olsen in Tennessee; McKernan in Georgia (Perissat, 1993). In 1989, Cuschieri and co-workers in Dundee performed the first laparoscopic cholecystectomy in the UK (Macintyre & Wilson, 1993). There followed a rapid, uncontrolled expansion of the procedure, described as '...the biggest unaudited free-for-all in the history of surgery' (Cuschieri, 1995). *Table 3* shows when the procedure was first introduced in a number of Western countries. Banta (1993) argued that surgeons adopted laparoscopic cholecystectomy under pressure from patients without knowing a great deal about its benefits and risks.

By 1992 laparoscopic cholecystectomy was fast becoming the procedure of choice in patients with symptomatic gallbladder disease (Zucker *et al,*

TABLE 3 Introduction of laparoscopic cholecystectomy in Western countries

Year of introduction	Country
1987	France
1988	USA
1989	Germany
1989	UK
1990	Australia
1990	Canada
1990	The Netherlands
1990	Sweden
1991	Denmark
Source: Banta (1993)	

1992). By 1995 it was estimated that this approach to cholecystectomy accounted for between 70% and 85% of such operations in the UK (Border, 1995).

A number of factors appear to be associated with the procedure's rapid diffusion. According to Macintyre and Wilson (1993), its popularity could be attributed largely to perceived benefits for the patient but there were also advantages for society, since less time off work was needed following the procedure (Macintyre & Wilson, 1993).

Sculpher (1993) argued that one key characteristic of laparoscopic cholecystectomy was the extent to which it had diffused widely as a result of perceived short-term benefits. He noted, however, that there had been little consideration of longer-term outcomes. The rapid adoption was also strongly influenced by a powerful combination of technological 'push' and the 'pull' of demand, according to Gelijns and Fendrick (1993), who described a variety of factors that influenced the widespread adoption of the procedure in the USA. These included provider competition, rapid regulatory approval of equipment, payers support of the new technique because it promised cost savings, and high patient demand because it promised to be less painful, to cause minimal scarring and to allow an earlier return to active life.

Gelijns and Fendrick (1993) pointed out that the diffusion of the technique in Europe occurred at only half the rate in the USA. They suggested a number of factors that might have contributed to

this, including differences in payment mechanisms. Banta (1993) also argued that the payment system was one of the most important factors in the diffusion of innovations such as laparoscopic cholecystectomy in European healthcare systems, in which many countries used budgeting systems aimed at limiting hospital expenditure. In the USA, payers such as Medicare were supportive of the new technique because it promised significant savings and American hospitals were reimbursed by Medicare at rates equal to those for conventional cholecystectomy (Gelijns & Fendrick, 1993). These authors also pointed out that European endoscope manufacturers, such as Storz and Wolf, were unable to meet European demands since they were initially focusing on the US market. In addition, European restrictions on the use of animals for training purposes might increase the time needed for a surgeon to achieve adequate clinical competence.

Pearson (1994) considered that laparoscopic cholecystectomy had been rapidly introduced without proper evaluation by an RCT of laparoscopic versus conventional cholecystectomy. RCTs of laparoscopic versus mini-cholecystectomy had been performed and had shown reduced length of stay, shorter convalescence and fewer complications for the laparoscopic technique. Improved cost-effectiveness of the technique over the conventional operation appeared likely because of the reduced length of stay but this had not been demonstrated unequivocally (Pearson, 1994).

According to Cuschieri (1995), the overriding reason for the unaudited expansion of MAS following the advent of laparoscopic cholecystectomy was the lack of effective central control in Western countries; many interventions in current practice had never been adequately evaluated.

By 1994, a Working Group on the implications of MAS for the NHS, commissioned by the Scottish Office Home and Health Department and the Department of Health, was forecasting that MAS would account for 70% of surgical procedures within 10 years (Cuschieri, 1994). According to Border (1995), even conservative estimates suggested that MAS approaches would account for 40% of all operations by the year 2000.

Potential benefits of MAS
The Working Group on the implications for the NHS of MAS stated that the confirmed benefits of MAS were reduced postoperative pain and ileus, accelerated recovery, lower incidence of postoperative respiratory complications, shorter hospital stay, early return to full activity or work,

and significant reductions in both early and late wound-related complications. They considered that other benefits were merely perceived and required confirmation; these included less immuno-suppression, and a decreased risk of adhesion formation and recurrent intestinal obstruction following surgical intervention (Cuschieri, 1994).

The likelihood that MAS procedures would allow patients to return to their usual activities faster than would be the case with conventional surgery has been a particularly strong 'selling-point' for the technology (Sculpher, 1993). A number of clinical papers on laparoscopic cholecystectomy (Reddick & Olsen, 1989; Barkun *et al*, 1992; Stoker *et al*, 1992) had focused much greater attention on the duration of convalescence than would normally be the case in clinical journals (Sculpher, 1993).

Potential disadvantages of MAS

Pearson (1994) lists the potential disadvantages of MAS as decreased quality of surgery, increased costs (of the operation), and the undertaking of oper-ations unnecessarily. Cuschieri (1995) noted that endoscopic surgery took longer and required more operating room time and facilities than conven-tional surgery. Further, its execution depended on novel technologies, which introduced a new vari-able in addition to traditional surgical skill. In his opinion, surgical performance during this type of surgery declined after about 4 hours. Another potential disadvantage of the new procedures was an enhanced risk of iatrogenic complications, directly via surgical complications and indirectly because of limited access; for example, following laparoscopic procedures for cancer, instances had been documented of tumour deposits in the access wounds for the operation (Cuschieri, 1995).

Sculpher (1993) drew attention to uncertainty about potential longer-term disadvantages of MAS caused by clinicians being unaware of outcomes, either because patients were not followed-up for long enough or because follow-up was not under-taken systematically. He queried whether patients were fully aware of the risks and uncertainties associated with longer-term outcomes when they agreed to undergo MAS procedures. Such proce-dures might involve fundamental changes to estab-lished surgical techniques, resulting in a complete redesign to cater for the MAS approach and the possibility of unforeseen adverse long-term outcomes (Border, 1995).

In addition to the trade-offs faced by patients when deciding whether to choose MAS or conventional surgery, there were other trade-offs to consider, such as that associated with the choice between MAS and non-surgical treatment (Sculpher, 1993). Non-surgical treat-ments include oral dissolution therapy, extra-corporeal shock wave lithotripsy and percutaneous approaches, although all of these are generally regarded as less effective than cholecystectomy (Macintyre & Wilson, 1993).

A potential disadvantage was that, in all large reported series of laparoscopic cholecystectomy, the operation could not be completed laparo-scopically in a proportion of patients and had to be converted to the open procedure; in emergencies, the rate tended to be higher (Macintyre & Wilson, 1993). These authors argued that surgeons could, however, ultimately expect to achieve a conversion rate of 2% or less.

Injury to the bile duct, resulting in a need for reconstruction, was the most serious complication associated with laparoscopic cholecystectomy, and one which commonly went unnoticed during the operation (Macintyre & Wilson, 1993). This major complication alone emphasised the need for adequate training and evaluation. Unrecognised injuries such as this carried high rates of morbidity and mortality, presumably because the early post-operative discharge caused a delay in presentation and diagnosis of the complication (Macintyre & Wilson, 1993).

Evaluation of laparoscopic cholecystectomy

Much of the literature on the evaluation of MAS procedures in general, including laparoscopic cholecystectomy, consists of observational studies (see *Table 4).* Three types of observational study used widely in the evaluation of healthcare tech-nologies are case series, cohort studies and case–control studies. These study designs are subject to various biases, which can undermine the quality of their findings, but they can usefully establish the potential effectiveness of a new technology and create a climate in which an RCT is seen as necessary in order to provide a definitive answer. Black (1996) argued that observational studies and RCTs should be viewed as complementary rather than competing methods of evaluation. In the case of MAS, Sculpher (1993) noted that the need for rapid generation of results to influ-ence clinical practice before widespread diffusion, the need for long-term follow-up and the potential difficulties in recruiting patients might mean that observational studies would be needed to supple-ment RCTs. It could be argued that this would be all that was possible.

TABLE 4 Laparoscopic cholecystectomy research categorised by study design

Study design	Number of studies 1987–March 1995 inclusive
RCT (experimental)	15
Non-randomised clinical trial	21
Retrospective cohort (observational and analytical)	19
Cohort with non-parallel comparison groups (observational and analytical)	21
Case–control study (observational and analytical)	0
Case series (of more than 100 cases) (observational and descriptive)	124
Source: Downs et al, *1996*	

Laparoscopic cholecystectomy has now been widely adopted and, although doubts have been expressed about the adequacy of the evaluations, the procedure has become established to the extent that the annual rate of publication in the medical literature relating to it has peaked and is now declining (see *Figure 11,* the data for which were obtained from a Medline search in June 1996, using the term 'laparoscopic cholecystectomy' in a combined MeSH and textword search).

Observational studies

Writing in 1994, 5 years after the first reported laparoscopic cholecystectomy, Pearson noted that the procedure had been assessed on the basis of a large number of prospective case studies but no RCT comparing laparoscopic with open cholecystectomy had been performed, so most claims of benefit were based on uncontrolled descriptive studies. Some of those benefits, such as reduced length of hospitalisation, shorter period of convalescence and fewer complications were, to some degree, supported by the results of RCTs of laparoscopic versus mini-cholecystectomy (Kunz *et al,* 1992; Barkun *et al,* 1992). However, it was no more than an informed guess that the same benefits applied when laparoscopic was compared with conventional cholecystectomy (Pearson, 1994). The results of the RCTs of laparoscopic versus mini-cholecystectomy are discussed below.

The early descriptive studies of laparoscopic cholecystectomy were summarised in a review by Macintyre and Wilson (1993); details of some of the larger studies are given in *Table 5.* Of the 28 studies, 22 were published in 1991 and six in 1992. The conclusions reached were that laparoscopic cholecystectomy had become the treatment of choice for patients presenting with gallbladder stones and patients could anticipate a hospital stay after their operation of less than 2 days. The incidence of bile-duct damage was higher than reported in case series of conventional
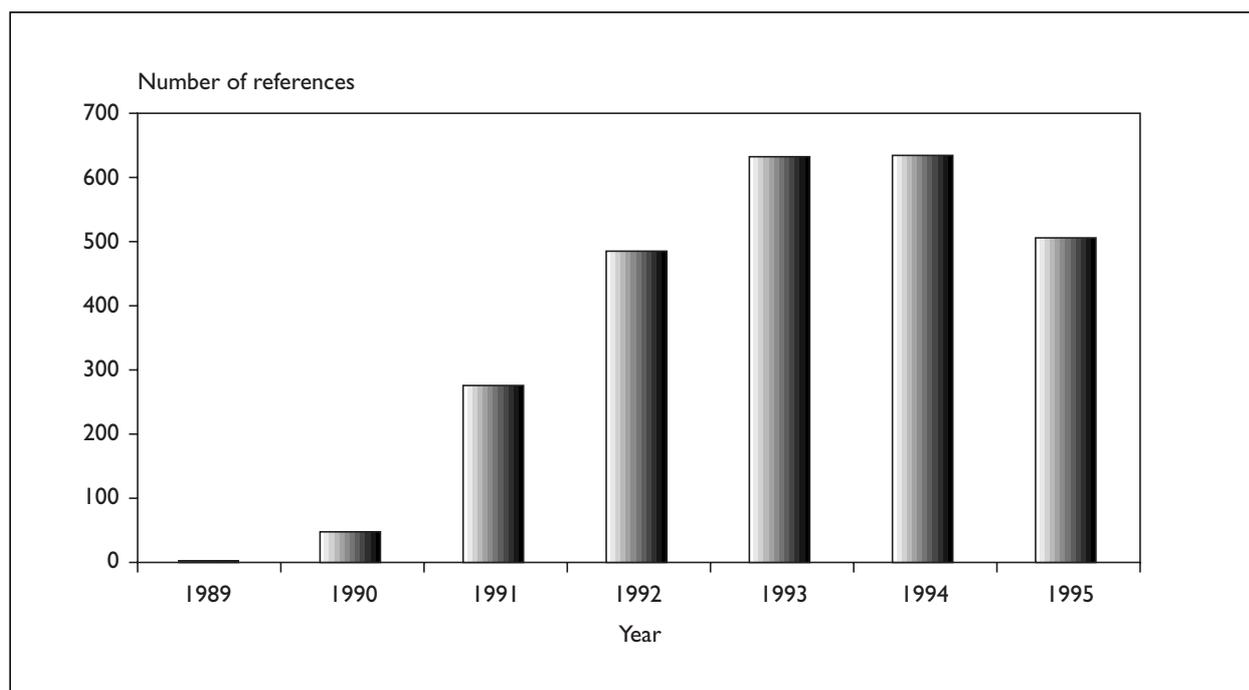
FIGURE 11 *Medline references – laparoscopic cholecystectomy 1989–1995*

TABLE 5 Selected observational studies – year of publication and patient numbers

Study	Year published	Patient numbers
1. Bailey *et al*	1991	375
2. Berci & Sackier	1991	418
3. Cuschieri *et al*	1991	1236
4. Dubois *et al*	1991	690
5. Ferzli & Kloss	1991	111
6. Goodman & Hunter	1991	100
7. Grace *et al*	1991	100
8. Graves *et al*	1991	304
9. McKernan	1991	50
10. Neugebauer *et al*	1991	100
11. Nottle *et al*	1991	50
12. Perissat *et al*	1991	104
13. Peters *et al*	1991	100
14. Schirmer *et al*	1991	152
15. Southern Surgeons Club	1991	1518
16. Spaw *et al*	1991	500
17. Voyles *et al*	1991	453
18. Walsh	1991	55
19. Western General Hospital, Edinburgh	1991	400
20. Wilson *et al*	1991	180
21. Wolfe *et al*	1991	381
22. Zucker *et al*	1991	100
23. Davis *et al*	1992	622
24. Graffis	1992	900
25. Larson *et al*	1992	1983
26. Martin *et al*	1992	162
27. Soper *et al*	1992	618
28. Stoker *et al*	1992	280

Source: Macintyre & Wilson (1993)

cholecystectomy. This was the major reported complication and remained a source of anxiety, although its occurrence could be reduced by good technique and appropriate training (Macintyre & Wilson, 1993).

Lack of a large-scale RCT of laparoscopic versus open cholecystectomy

The RCT is generally held to be by far the most satisfactory method of clinical assessment in most circumstances. Randomly allocating patients between treatments minimises selection bias by balancing not just known prognostic factors but also unrecognised and unmeasurable factors (Department of Health, 1992). Systematic errors can be avoided by ensuring that random allocation is followed by an unbiased statistical analysis that includes all those who were randomised. If randomisation is not used, the effects of moderate biases can either obscure moderate but worthwhile effects, or give an impression of benefits when they do not exist (Department of Health, 1992).

In 1989, before widespread diffusion, Cuschieri stated that prospective RCTs were needed to define the indications for the laparoscopic approach and to confirm its benefits against the standard cholecystectomy, which would still be needed in a proportion of patients. McMahon and colleagues (1992) reported on a survey of surgeons and research ethics committees undertaken in 1992 on the necessity and ethics of an RCT to compare laparoscopic with open cholecystectomy. This showed wide support for a trial comparing the techniques, although respondents with more experience in laparoscopic cholecystectomy were less convinced of the need for a trial. However, the rapidly increasing demand by professionals and patients for a less invasive way to remove gallstones led to laparoscopic cholecystectomy being adopted by surgeons before an RCT was carried out (Gelijns & Fendrick, 1993). This demand for the new method, and the accompanying resistance to using the old, was seen as one of the major problems in evaluating the procedure (Pearson, 1994).

Downs and colleagues (1996) identified ten small RCTs of laparoscopic and open cholecystectomy (see *Table 6* for patient numbers), one of which (Coelho *et al*, 1993) compared laparoscopic with both open and mini-cholecystectomy. General findings in respect of the laparoscopic procedure included longer operation time, less postoperative pain and a shorter postoperative hospital stay.

TABLE 6 RCTs comparing laparoscopic with open cholecystectomy

Study	Patient numbers
1. Putensen-Himmer *et al*, 1992	20
2. Agnifili *et al*, 1993	50
3. Coelho *et al*, 1993	45
4. Garcia-Caballero & Vara-Thorbeck, 1993	100
5. Jan & Chen, 1993	101
6. McMahon *et al*, 1993a	63
7. Schauer *et al*, 1993	40
8. Trondsen *et al*, 1993	72
9. Berggren *et al*, 1994	30
10. Byrne *et al*, 1994	32
Source: Downs et al, *1996*	

TABLE 7 RCTs of laparoscopic versus mini-cholecystectomy

Study	Patient numbers
1. Barkun *et al*, 1992	70
2. Kunz *et al*, 1992	77
3. Coelho *et al*, 1993	45
4. McMahon *et al*, 1993b	20
5. Tate *et al*, 1993	22
6. McMahon *et al*, 1994a	299
7. McMahon *et al*, 1994b	133
8. McGinn *et al*, 1995	300
9. McMahon *et al*, 1995	299
10. Peacock *et al*, 1995	60
11. Majeed *et al*, 1996	193
Source: Downs et al, *1996*	

RCTs of laparoscopic versus mini-cholecystectomy

Details of RCTs of laparoscopic versus mini-cholecystectomy, including patient numbers, are presented in *Table 7.*

The results of early trials were very favourable towards laparoscopic cholecystectomy. (Kunz *et al,* 1992; Barkun *et al,* 1992; McMahon *et al,* 1994a;b). One trial (Kunz *et al,* 1992), from October 1990 to October 1991, showed that the advantages included less postoperative pain, less restriction of total vital capacity, and a shorter postoperative hospital stay. Another trial (Barkun *et al,* 1992), from September 1990 to September 1991, also suggested that laparo-scopic cholecystectomy was preferable, since it reduced the length of hospital stay, the duration of convalescence and the number of complications. McMahon and colleagues (1994a;b), reporting a trial covering August 1991–March 1993, found a more rapid return to work, better physical and social functioning, and less pain and depression after the laparoscopic procedure. These differences became smaller over time, however, and by 3 months postoperatively there were no differences, except that the laparoscopy patients were more satisfied with the appearance of their scars.

McMahon and colleagues (1995) followed-up their patients with a symptoms questionnaire a year after they had taken part in the original trial. They found that overall, the laparoscopic method did not have any clear symptomatic advantage over mini-cholecystectomy, the only difference being that significantly fewer laparoscopy patients reported heartburn.

The results of two more recently published RCTs comparing laparoscopic with mini-cholecystectomy have been less favourable toward laparoscopic cholecystectomy. McGinn and colleagues (1995), in a trial covering June 1991–April 1995, reported a longer operating time, a greater conversion rate and a greater complication rate for the laparoscopic proce-dure but a quicker return to normal activities; they concluded that no clear advantage for either operation had been demonstrated. McGinn and colleagues considered that new techniques and operations, like new drugs, ought to be evaluated by proper RCTs, despite claims when this trial was initiated, that it was neither practical nor ethically possible.

Majeed and colleagues (1996), in a trial covering January 1992–June 1995, found that the laparo-scopic procedure took longer, with no significant advantage in terms of hospital stay or postoperative recovery. One reason why these findings are less favourable may lie in the trial design. Patients and carers were successfully blinded in the immediate postoperative period: identical wound dressings were applied; great care was taken to relieve pain; patients themselves determined the time of dis-charge; no advice was issued on the time needed for convalescence (Terpstra, 1996).

Downs and colleagues (1996), in a systematic review of the effectiveness and safety of laparoscopic cholecystectomy, identified 11 reports of RCTs (some of which referred to the same trial) comparing laparoscopic and mini- or open cholecystectomy, including one that compared laparoscopic with mini- and open cholecystectomy. Of these 11 reports, two were published in 1992, three in 1993, two in 1994, three in 1995, and one in 1996. On the basis of this, Downs and colleagues (1996) made the following recommendations.

1. Surgeons should not be encouraged to replace mini-cholecystectomy with laparoscopic.
2. A valid system for classifying and grading complications by their severity should be promoted.
3. A study is required with standardised data collection, long and defined follow-up and high external validity (all operations on a geographically defined population) to:
   (a) compare complication rates (particularly long-term complications such as biliary stricture) between open, mini- and laparoscopic cholecystectomy
   (b) assess the impact of surgeons' training, overall experience, and the frequency with which they undertake mini-, open and laparoscopic cholecystectomy, on postoperative outcome
   (c) assess the effectiveness of cholecystectomy as a treatment for abdominal pain, including the frequency of symptom recurrence, quality of life and patient satisfaction.

Barriers to adequate evaluation of laparoscopic cholecystectomy

The evaluation of laparoscopic cholecystectomy posed several quite diverse major problems, some of which apply to MAS in general. They include varying skill levels, rapid technological change, patient preferences, effective blinding, long-term costs and ethical questions.

Varying skills

There were difficulties associated with varying levels of surgical skill. For example, a shortfall of skill in those at an early point on the learning curve might result in poor performance by operators. Sculpher (1993) drew attention to some of the implications of the learning curve for the use of RCTs. One was that it was likely that RCTs would be undertaken by clinical enthusiasts with high levels of skill who were often based in medical schools. If the clinicians who conducted the trial were not representative of the practitioners who

would undertake the procedure following its diffusion, then the generalisability of the RCT might be questionable. The problem might be overcome, however, by undertaking the trial at a number of centres that were more generally representative of clinical practice (Sculpher, 1993).

Rapid technological change

Another rather different problem was the potential speed of change of MAS technologies, which was partly a function of the parallel development of similar technology in non-medical applications. A feature of the 'assessment' part of HTA, as defined by the NHS Executive (1996), was that the health technology should be reasonably stable and be able to be compared with competing health technologies and/or no intervention. The value of the results of an RCT would be diminished if, midway through a study, the clinicians considered that the procedure should be altered to reflect new and important developments (Sculpher, 1993).

Patient preferences

Yet another difficulty encountered when RCTs were used to evaluate MAS procedures was patient recruitment. Given the apparent reduction in trauma and length of convalescence associated with laparoscopic procedures, Sculpher found it difficult to persuade patients to enter a clinical trial where they had a 50% chance of being randomised to the alternative form of therapy. Arguably, this suggests that patients were not being fully informed about uncertainties, about possible risks or long-term problems. Paradoxically, it is in such situations of patient-led demand, where MAS procedures are diffusing rapidly, that there is a real need to undertake RCTs, to clarify the risks and benefits, and to disseminate their results as quickly as possible (Sculpher, 1993).

Media reports

There is general agreement that media reports of the apparent benefits of the laparoscopic method led to patients becoming aware of the procedure and requesting it in preference to open cholecystectomy. Banta (1993) argued that the most important force facilitating the diffusion of MAS, such as laparoscopic cholecystectomy, was patient demand which, along with physician interest, was fostered by press reporting. White (1993) noted that reports in the lay press detailed the advantages of a marked reduction in recovery time and postoperative pain and a better cosmetic result, and that patient demand for the procedure soon became overwhelming. These reports will have emanated from those using the operation and suggest that

surgeons should have been more cautious in their claims for it. It is therefore ironic that, in addition to reporting the perceived benefits of laparoscopic cholecystectomy, the media also reported instances where complications arose (see, for example, *The Times,* 8 December 1992). McGinn and colleagues (1995) noted that although none of their patients refused to enter the trial, during the first 1–2 years of their RCT it was sometimes difficult to persuade patients to take part if they had been referred specifically for a laparoscopic procedure; however, recently the opposite trend had been noted following adverse media publicity about the technique.

Effective blinding

Pearson (1994) pointed out the difficulty of trying to achieve blinding in any RCT of a MAS procedure versus the conventional operation: the nature of the scars would be obvious to both patient and assessor, and it would be difficult to correct for the effect of this knowledge. Terpstra (1996) noted that blinding of patients and carers was successfully achieved in the RCT comparing laparoscopic and mini-cholecystectomy undertaken by Majeed and colleagues (1996). It can be argued, however, that blinding does not test the operation as it would actually be used. The sight of a small scar may, for example, help the patient to recover more quickly, and this is an alternative explanation for the different findings in the Majeed trial.

Long-term costs and outcomes

Clearly the long-term implications of MAS procedures also needed to be taken into account when considering how to evaluate these technologies. RCTs would need large patient numbers and long follow-up periods in order to detect important but rare adverse events. The cost and logistical difficulties involved have generally inhibited such trials (Sculpher, 1993). McMahon (1995) pointed out that there are no controlled data comparing the long-term outcomes of laparoscopic and open cholecystectomy.

Ethical questions

As early as 1993, Macintyre and Wilson stated that the opportunity to conduct a large-scale controlled trial comparing open and laparoscopic cholecystectomy had now passed and suggested a number of reasons for this. Simple audit rapidly demonstrated the shorter hospital stay and reduced time off work associated with the laparoscopic procedure. Surgeons themselves were impressed with patients' feelings of well-being and relative lack of pain. In such circumstances, a clinical trial was considered unethical by most surgeons, despite the lack of reliable evidence about rare or longer-term outcomes.

A large multi-centre group in Germany did consider mounting a controlled trial (Neuge-bauer *et al,* 1991) but decided to postpone it while surgeons learned to use the new procedure safely and effectively. The trial never took place, however, because by the time the first 100 patients had been operated on it was being argued that the obvious benefits of the laparoscopic procedure (quicker recovery, less pain, reduced length of hospital-isation) placed ethical constraints on undertaking an RCT.

The ethical dimension of randomisation in relation to equipoise, where there is no preference between the treatment options, has been discussed by Lil-ford and Jackson (1995). The 1994 MRC report stated that in deciding whether a particular patient should be entered into an RCT, the clinician and patient should both be uncertain about the relative merits of the alternative treatments being offered.

Baxter and O'Dwyer (1992) argued that it was dangerous, if not unethical, to accept any new treatment as being significantly better than an existing one that gives excellent results, without putting it to the ultimate test of an RCT. McMahon and colleagues (1992) reported on a survey of 40 ethics committees who were sent a hypothetical protocol for a trial comparing laparoscopic with open cholecystectomy: 25 gave approval, 12 refused to comment, and only three considered such a trial unethical. Thus, the balance of opinion in these committees at that time favoured Baxter and O'Dwyer's viewpoint.

The experience of laparoscopic cholecystectomy suggests that surgeons and their patients should have remained uncertain about whether to use the technique for much longer, until reliable evidence became available.

Economic aspects

Sculpher (1993) argued that two critical economic issues needed to be formally addressed before new technologies diffused widely within the health service. Assessments were needed to find out how the costs of new technologies compared with those of existing ones and to identify what additional benefits, in terms of patient health outcomes, were generated by these new technologies.

Capital and operating costs

Border (1995) stated that MAS might reduce costs through shorter hospital stays and periods

of convalescence, and by reducing complication rates and analgesic requirements, while wider economic benefits might be gained through earlier return to work. However, longer operating times, high capital and running costs, and the possibility of an increase in both medical complications and demand for operations might all act to increase costs. The Working Group on the implications of MAS for the NHS estimated the capital costs involved in setting up an operating theatre for MAS as roughly £30,000. Recurring costs would be influenced by the level of use of disposable instruments, setting-up times for MAS operations, operating times, and length of stay (Cuschieri, 1994).

Macintyre and Wilson (1993) argued that any economic assessment should also take into account the benefits to society of an earlier return to work. They noted that this was a complex area, with the little published material available providing conflicting conclusions, and cited studies by Kurzawinski and colleagues (1992), Hardy and colleagues (1992), and Macintyre and colleagues (1992) as demonstrating lower hospital costs for laparoscopic than for open cholecystectomy, while a study by Stoker and colleagues (1992) showed the opposite. Macintyre and Wilson (1993) maintained that the resource benefits to the healthcare system as a whole were difficult to assess and were not confined to the results of shorter postoperative stays. Shorter stays should result in reduced bed requirements and reduced costs for drugs and disposables but there was the capital cost of instruments and equipment, including depreciation and maintenance. The use of disposable equipment further increased costs. There were also costs associated with moving clinicians up the learning curve, surgeons with less experienced being more likely to encounter complications (Macintyre & Wilson, 1993).

According to Pearson (1994), laparoscopic might be more cost-effective than conventional cholecystectomy but this had yet to be demonstrated. If it was more cost-effective, it would be mainly because of the decreased length of hospital stay and reduced morbidity compared with open and mini-cholecystectomy.

Bass and colleagues (1993) estimated the overall cost-effectiveness of laparoscopic compared with open cholecystectomy in terms of expected hospital charges and quality-adjusted months of life over a 5-year period. These estimates were incorporated into a computer model designed:

- to compare expected short-term and long-term outcomes of both procedures

- to compare expected direct costs associated with each treatment
- to determine how the cost-effectiveness of laparoscopic relative to open cholecystectomy varies as a function of patient age and gender.

Bass and colleagues (1993) concluded that laparoscopic was likely to be less costly and more effective than open cholecystectomy for most patients, as long as it did not routinely require pre-operative cholangiography and was not associated with increased professional fees or increased risks of retained stones or bile-duct injury. They quoted figures for 45-year-old women and men for laparoscopic ($5354; $6036) and open cholecystectomy ($5525; $6830) that represented the projected total average charge after 5 years. One-way and multi-way sensitivity analyses were performed to assess how expected charges and survival varied according to patient age and sex, and with the varying estimates of the probability, utility, and charge parameters. The difference in favour of laparoscopic cholecystectomy rose substantially with increasing age for both women and men (Bass *et al,* 1993).

Kesteloot and Penninckx (1993) undertook a study of the costs and effects of open versus laparoscopic cholecystectomy, based on a consecutive series of 47 patients who underwent cholecystectomy between November 1990 and February 1991 in the University Hospital Gasthuisberg in Belgium: 21 patients underwent laparoscopic cholecystectomy and 26 underwent the open procedure. The difference in hospital costs between the two procedures depended on the operating theatre equipment, time spent in the operating theatre, materials used during the procedure, variations in postoperative length of stay, and use of materials during each postoperative patient day. Kesteloot and Penninckx concluded that, with more experience, most hospitals could realise cost-savings by switching, as far as was medically justified, to laparoscopic procedures.

A study by Fullarton (1992) found that laparoscopic cholecystectomy at £2053 was cheaper than open cholecystectomy at £2250, and Kurzawinski and colleagues (1992) reported figures for the laparoscopic procedure of £1938 (£2144 if disposable instruments were used) and £2172 for open cholecystectomy. The Working Group on the implications for the NHS of MAS quoted figures of £1206 for laparoscopic cholecystectomy(£1509 if disposable instruments were used) and £1114 for mini-cholecystectomy (Cuschieri, 1994).

Few of the RCTs identified contained any economic analysis. Jan and Chen (1992), in a study involving

101 patients, found that laparoscopic was more expensive that open cholecystectomy. McMahon and colleagues (1994a;b), in a study involving 302 patients, reported that laparoscopic, at £1183 (£1486 if disposable instruments were used), was more expensive than mini-cholecystectomy, at £1090. Costs at the two main hospitals involved were used to estimate average theatre and ward costs. The NHS cost of hospital admission for each patient was estimated from recorded time in theatre, hospitalisation, and additional tests or treatment. The mean total cost per patient was £396 greater for the laparoscopic method (95% confidence interval (CI) £328–£465). If disposable instruments had not been used, the difference in mean costs would have been £93 (95% CI £25–£162). Operating time (and therefore theatre cost) was greater for laparoscopic cholecystectomy, while hospital stay (and therefore ward cost) was greater for mini-cholecystectomy.

Barkun and colleagues (1995), in their 1991–92 trial, found that laparoscopic at $1908 per patient was cheaper than mini-cholecystectomy at $2106, and was associated with improvements in the rate of postoperative convalescence, confirming the cost advantages of the laparoscopic approach in this setting. The cost analysis was based on data from 68 patients, and its purpose was to quantify the total direct costs to the healthcare system of each strategy of management; no specific allowance was made for the additional cost of training or personnel to operate the specialised instrumentation required by the laparoscopic procedure or for the costs of possible complications, such as a common hepatic duct injury. Costs were recorded for the preoperative phase, the hospital phase and the first 18 months of follow-up (Barkun *et al,* 1995).

McGinn (1995) gave no figures but reported no significant differences between laparoscopic and mini-cholecystectomy costs.

The estimated costs from four of the above studies of laparoscopic (using re-usable instruments) compared with mini- or open cholecystectomy are shown in *Figure 12.* Laparoscopic was found to be more expensive than open but less expensive than mini-cholecystectomy.

These reports do not support any general conclusions about the relative costs of the procedures. There are, of course, many reasons for variations in costs between sites; for example, one such cost is specifically a function of the experimental situation. As Sculpher (1993) stated, one concern about RCTs was that trial protocols might impose atypical patterns of care on unrepresentative samples of patients which would make the observed resource use in the trial difficult to generalise into routine clinical practice.
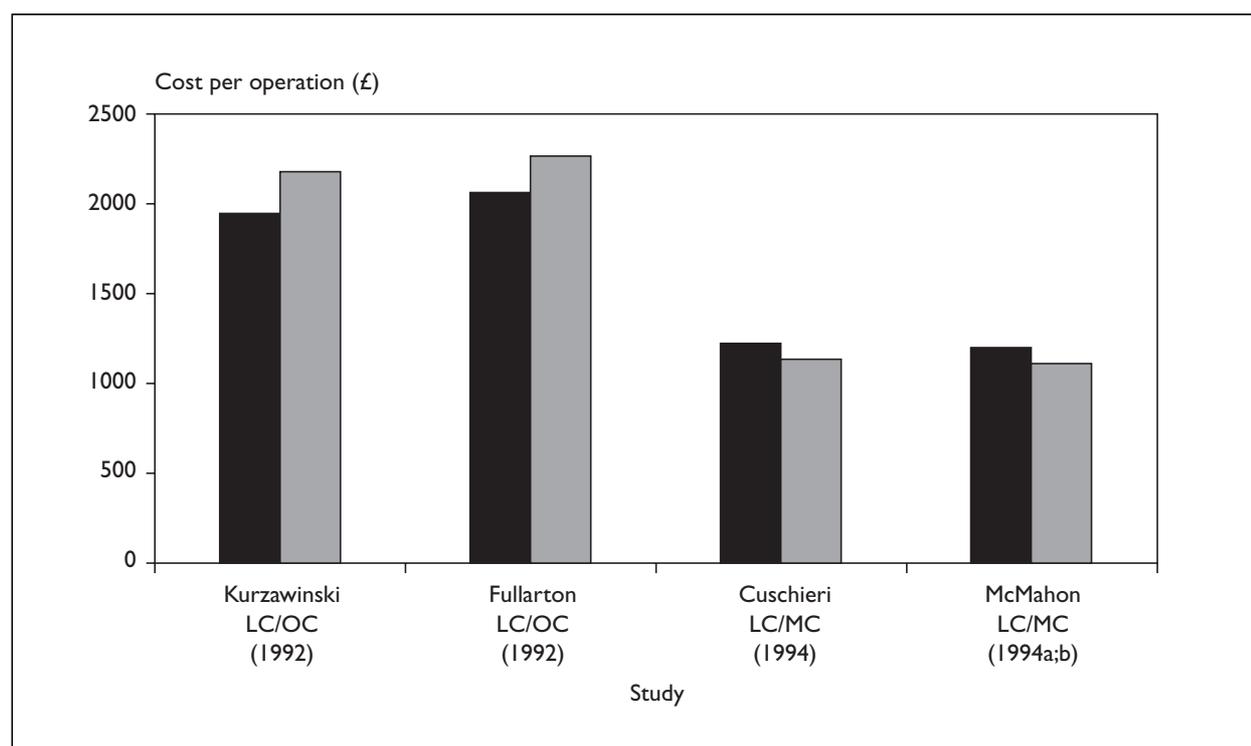
FIGURE 12 *Relative costs of laparoscopic cholecystectomy (LC), mini- (MC) and conventional cholecystectomy (OC) (■, LC; □, MC/OC)*

Sculpher believed that trial data would need to be supplemented with outside data using models that synthesised information from a variety of sources. Another potential problem of using RCTs to collect resource data would arise if the clinical uncertainty became resolved. Achieving a sufficiently large sample size to calculate resource use would depend on the trial continuing (Sculpher, 1993) but was unlikely to happen once clinical criteria had been met.

Training costs

It has been acknowledged that MAS, including laparoscopic cholecystectomy, requires special skills. Training costs for the adoption of new medical technology are often high (Pearson, 1994); this is certainly the case for MAS and is, therefore, part of the cost of introducing MAS. In the USA, such costs are largely borne by industry, which built commercial training centres for MAS techniques to provide surgeons with hands-on experience and to introduce them to procedure-related products (Gelijns & Fendrick, 1993). More than half of the 32,750 practising surgeons in the USA received training in laparoscopic cholecystectomy during the 18 months after the procedure was introduced (Gelijns & Fendrick, 1993). At that time, UK surgeons did not have the same training opportunities and Cuschieri (1989) considered that there was genuine concern and a real risk that the procedure would be taken up in the UK by surgeons without proper training.

The implications of differences in training between the USA and the UK were discussed by Pearson (1994). In the USA, the 'credentialling' process specified, in a written document, those operations that a surgeon was fully trained to perform. Employers then granted 'privileges' to surgeons that stated which operations they were allowed to perform. The nearest equivalent to this in the UK was accreditation for Higher Specialist Training. This was awarded on completion of training but did not specifically list the skills acquired. The differences between the two systems were important; surgeons in the UK who undertook MAS after a preliminary course with no additional training were significantly more likely to experience complications during surgery than those who had undergone additional training (Pearson, 1994).

Macintyre and Wilson (1993) reported that the English and Australasian Surgical Colleges, and the Society of American Gastrointestinal Endoscopic Surgeons had all made recommendations on training for laparoscopic cholecystectomy. In

essence these were that the surgeon performing the procedure should be a trained general surgeon with biliary experience, with a knowledge of diagnostic laparoscopy, who should have attended a course or workshop – preferably one which offered hands-on experience. The English College also recommended visiting a centre where the technique was already established. A UK survey of laparoscopic cholecystectomy practice, undertaken by Macintyre and Wilson (1993), revealed that only 40% of consultants performing cholecystectomy had previously performed laparoscopy on a regular basis.

Reports in the media in the early 1990s of patients sustaining serious injuries as a result of MAS procedures helped to focus on the lack of formal training and, in 1993, the UK Government made funds available to establish special MAS training centres. Since then four centres have been established, in London, Leeds, Dundee and Cardiff. Their role is more to provide facilities for trainees to develop their skills by practising on artificial organs/tissues and organs removed from dead animals, rather than to involve them in supervised assistance with operations on real patients (Border, 1995).

In 1994, the Senate of the Royal Surgical Colleges published new training requirements covering not just MAS but all new and existing techniques across surgical specialities (Border, 1995). Cuschieri (1995) recommended that endoscopic training skills should only be taught within an academic environment by dedicated experts supported by a team of full-time technician tutors. He saw the emergence of the regional endoscopic training centres as a logical development in MAS (Cuschieri, 1995).

The Working Group on the implications of MAS for the NHS set out a number of practical considerations for safe MAS, including training recommendations (Cuschieri, 1994). These are reproduced in Appendix 3.

Appropriate allowance for the considerable costs of these training arrangements must be included in any calculation of true economic aspects of laparoscopic cholecystectomy. Training costs were not included in a number of studies (Bass, 1993; McMahon *et al*, 1994; Barkun *et al*, 1995; Kesteloot & Penninckx, 1993).

Effect of laparoscopic cholecystectomy on cholecystectomy rates

Unit costs of laparoscopic cholecystectomy are not the only consideration. There is also the effect on

the pattern of demand for cholecystectomy following introduction of a procedure which is more attractive to patients. Cuschieri (1995) drew attention to the phenomenon whereby new advances create an increased workload and thereby raise overall expenditure. He pointed out that the cholecystectomy rate had increased in Scotland and worldwide since the advent of laparoscopic cholecystectomy (see *Table 8* for Scottish data). He suggested a number of possible reasons for this. First, patients with low-grade symptoms perceived the new surgery as more acceptable and were now willing to undergo a surgical treatment which inflicted less postoperative pain and resulted in a shorter period of disability. Second, surgeons' perception of the risk–benefit ratio was altered, resulting in a change in the indications for cholecystectomy. Third, gastroenterologists might increase referrals for surgery because they viewed laparoscopic cholecystectomy as more effective than alternative non-surgical treatments (Cuschieri, 1995). An increase of almost 20% in the numbers of cholecystectomies undertaken in Scotland during the period 1990–95 is shown in *Table 8*, which demonstrates that the procedure has come to be seen as the preferred option.

TABLE 8 Open and laparoscopic cholecystectomies undertaken in Scotland

Year	Open cholecystectomy	Laparoscopic cholecystectomy	Total
1990	4354	107	4461
1991	3257	1315	4572
1992	2431	2405	4836
1993	1940	3390	5330
1994	2025	3263	5288
1995	1931	3418	5349
Source: Information and Statistics Division, NHS in Scotland			

Bloor and Maynard (1994) argued, however, that hospital managers were increasingly questioning the appropriateness of procedures such as laparoscopic cholecystectomy, which required the purchase of sophisticated and expensive capital equipment and increased theatre time, thus resulting in lower patient throughput for surgical procedures. They pointed out that it was estimated that MAS procedures such as this would continue to take about one-third longer than conventional operations, with complications increasing this difference in time.

A sharp rise in the rate of cholecystectomies since the introduction of laparoscopic procedures in the USA was documented by Legorreta and colleagues (1993). Although the rate of open procedures declined between 1988 and 1992, there was an increase in the total number of cholecystectomies leading to a 17.8% increase in Health Maintenance Organisation medical expenditure for cholecystectomy over this period, despite a reduction of 25.1% in unit costs. These findings contrasted with the generally stable cholecystectomy rates in the USA during most of the 1980s.

Macintyre and Wilson (1993) stated that although the only absolute contraindication to open cholecystectomy was unfitness for general anaesthesia, the laparoscopic procedure had several additional contraindications. As the operation has developed, however, these have diminished to three: unfitness for general anaesthesia; cholecystoenteric fistula; and doubts about possible gall bladder malignancy.

How should surgical innovations be evaluated?

This review of the major reported evaluations of laparoscopic cholecystectomy has revealed many of the problems that arose in the absence of a systematic approach to evaluation. Recently, a number of recommendations have been made for a more systematic approach to the evaluation of new surgical techniques, in general, and MAS and laparoscopic cholecystectomy, in particular.

Cuschieri (1995) argued that a window of opportunity followed a new medical and surgical advance, during which prospective randomised studies were possible and feasible. Once the procedure became widespread, the window closed and recruitment of patients for RCTs became impossible. This suggests that evaluation should be initiated at a very early stage. In addition, valid comparisons between two surgical options depended on proficiency with both procedures but, by the time surgeons had acquired the necessary familiarity with the new approach, they had often come to prefer it, for a variety of reasons, and were thus unable or unwilling to participate in trials (Cuschieri, 1995). The implication is that it may prove extremely difficult to get an adequate number of truly randomised trials even when starting early. Banta (1993) stated that the effectiveness of any new procedure needed to be established by well-controlled clinical trials and that the evaluation of effectiveness needed to be part of the diffusion process. Thus, procedures for on-going evaluations that address the problems of the difficulty of randomisation after the early stages of a trial need to be agreed upon.

Minimal access surgery

A number of writers have made more detailed suggestions on how surgical innovations should be evaluated. Bloor and Maynard (1994) argued that pioneering innovations such as MAS should and could be carefully evaluated in well-designed trials. They pointed out that although there were difficulties in implementing RCTs of surgical techniques due to the lack of blinding, a carefully designed trial could overcome those problems. They recommended that, ideally, an economic evaluation should be carried out in parallel with the clinical trial. Once the effectiveness and efficiency of a new technique was established, it should be introduced gradually, with careful training of surgeons, and accreditation procedures in place. Gross (1993) also argued that any new surgical procedure competing with current practice should be subjected to an RCT before it was available for general use, although Bouchard and colleagues (1996) argued that not all new technologies could or should undergo a formal randomised comparative assessment.

Pearson (1994) held that MAS techniques should be assessed as part of a research programme to ensure that surgeon training and clinical outcomes were evaluated. There was a danger that, unless the introduction of MAS techniques was strictly controlled by purchasers, they would become part of routine clinical practice before their effectiveness and cost-effectiveness had been fully assessed. Purchasers are recommended to consider the implications of new techniques for service planning and use the contracting process to prevent the adoption of unevaluated techniques which they did not wish to purchase. In addition, purchasers should require routine monitoring/audit to be undertaken for those techniques that were introduced (Pearson, 1994).

Border (1995) described a number of alternatives which had been suggested for evaluating MAS. These included:

- pragmatic clinical trials, involving surgeons at different points on the learning curve
- monitoring of new techniques at selected centres and their performance compared against known standards for established methods
- a combination of approaches, possibly starting with laboratory or animal-based studies, moving on to prospective evaluations, then to full-scale clinical trials, followed by a surveillance study involving long-term follow-up of a selected population.

Reviewing the economics of MAS, Sculpher (1993) noted that the clinical and economic evaluation of many forms of MAS remained haphazard. He drew attention to the fact that some procedures were likely to diffuse, despite being less cost-effective than the conventional treatment, because of factors such as a lack of systematic evaluation, or because the results of evaluations were poorly disseminated or ignored by practitioners. For the same reasons, other MAS procedures, which were cost-effective, were likely to remain under-utilised. In order to address issues such as these, Sculpher (1993) argued that policies designed to achieve the rational diffusion of MAS should seek to achieve the following objectives.

1. At an early stage in their development, MAS applications with a clear potential to improve patient care should be identified (according to explicit criteria such as being likely to reduce costs or to improve benefits and being relevant to large numbers of patients).
2. Adequate funds should be ensured for a programme of well-designed and timely clinical and economic evaluations of these applications culminating, where possible, in appropriately powered pragmatic RCTs.
3. Given that perfect information will not necessarily ensure rational behaviour, barriers and incentives should be provided, so that the speed and extent of diffusion of new MAS procedures closely reflects the results of the evaluations.
4. Regulations should be introduced to ensure that practitioners have sufficient training to undertake procedures and to achieve the sorts of results observed in controlled trials.

According to Cuschieri (1995), the way in which MAS procedures should be evaluated was dependent on the category of the operation. He argued that the category was 'equivalent' if the nature and steps of the new operation were unchanged from the validated procedure with which it was being compared. He maintained that within this category, into which he placed laparoscopic cholecystectomy, RCTs were unnecessary. Instead, prospective audit of outcome studies should provide the necessary information for valid assessment. However, if the MAS procedure was new or incorporated significant changes from the conventional counterpart, then Phase II studies in key centres should constitute the initial phase of evaluation. If these suggested benefit then RCTs to compare the two approaches were warranted (Cuschieri 1995).

Advisory Council on Science and Technology

Official committees and professional bodies have also addressed the task of producing guidelines.

In the UK and elsewhere, detailed recommendations have been made for the evaluation and control of new surgical techniques. The report of ACOST on medical research and health sought to identify how advances in science and technology could be used to provide better health cost-effectively (ACOST, 1993). A number of task forces were set up by ACOST, one of which focused on surgical technologies. With specific reference to MAS procedures, the task force provided a number of policy suggestions regarding surgical technologies, including the following.

1. All new medical devices or novel applications of existing devices to be used within the NHS should be developed only under controlled conditions.
2. Only procedures and equipment that have undergone assessment and approval should be used within the NHS.
3. Novel surgical procedures and surgical teams should be registered with a Committee on the Safety and Efficacy of Procedures.
4. The NHS should follow codes of practice in deciding the most cost-effective way of introducing novel devices/applications.
5. Specific centres specialising in appropriate diseases and techniques should be adequately researched to develop, to evaluate and to educate the rest of the profession.

The government, however, rejected the first two recommendations above. According to Border (1995), although the Department of Health recognised the need for stricter controls on the introduction of new surgical procedures, it viewed regulation in this area as a matter for professional judgement and guidance rather than legislation.

Senate of the Royal Surgical Colleges of Great Britain and Ireland

Professional bodies have also proposed procedural rules for evaluating new techniques. Border (1995) described a possible scheme for a control system for surgical techniques proposed by the Senate of the Royal Surgical Colleges of Great Britain and Ireland. The proposal was designed to operate in stages. Once a new development had been identified, its use would be restricted initially to specified centres, where it would be evaluated in clinical trials. When the value of the new technique in terms of its safety, effectiveness and cost-effectiveness had been established, its use would be restricted to those surgeons who had received appropriate training. Eventually, if the procedure became sufficiently widely used, it would be included as part of the basic surgical skills

assessment. Finally, the performance of the new procedure and of the surgeons using it would be monitored by a process of on-going clinical audit (Border, 1995).

SERNIP

The most recent developments have arisen out of discussions between the Department of Health and the Royal Colleges on a new system to bring together groups of experts to evaluate major advances in surgery and medicine. This has led to the establishment of SERNIP, a voluntary register of new procedures, which aims to protect patients from the inappropriate application of new interventional procedures whose safety and efficacy have not been established (Department of Health, 1995).

SERNIP incorporates a method of identifying and registering such procedures and advising on how they may be evaluated in a controlled manner. Those new procedures whose safety and efficacy have not yet been established and which SERNIP classifies as requiring a fully controlled evaluation may only be used as part of systematic research, comprising either an observational study, in which all interventions and their outcomes are systematically recorded, or an RCT (Department of Health, 1995).

One of the difficulties in all of these proposed regulatory systems is the lack of precise interpretation of the term 'new'. While radical departures from accepted procedure are easy to identify, there remains the question of how much can a procedure be modified before it becomes a 'new' procedure, subject to an independent evaluation? Another problem is deciding what constitutes evidence of 'efficacy'.

Discussion

The publication trends displayed in *Figure 11* showed that the annual numbers of papers relating to laparoscopic cholecystectomy peaked in 1993–94 and began to decline in 1995, thus indicating that this procedure is widely accepted by surgeons as acceptable and effective for well-specified indications. However, the papers reviewed above suggest that the case for this acceptance is not wholly convincing. This review shows that evaluation did not begin until the technology had diffused quite widely, by which time it was no longer seen as possible to carry out fully controlled and satisfactory RCTs of laparoscopic versus open cholecystectomy. Observational

studies of laparoscopic cholecystectomy reported benefits of a shorter hospital stay, less post-operative pain, more rapid return to full activity and a smaller scar, although there was also, at least initially, an increased likelihood of injury to the bile duct. Laparoscopic cholecystectomy was also found to be less costly than open cholecystectomy. The early RCTs comparing laparoscopic and mini-cholecystectomy also found in favour of the laparoscopic procedure for much the same reasons, although economic analysis, where included, indicated that it was more expensive than mini-cholecystectomy. However, two later RCTs concluded that the laparoscopic procedure held no significant advantages. One conclusion of Downs and colleagues (1996), in their systematic review of the effectiveness and safety of the procedure, was that surgeons should not be encouraged to replace mini- with laparoscopic cholecystectomy.

Even the combined results of the observational assessments, the limited RCTs of laparoscopic versus open cholecystectomy and the more extensive RCTs of laparoscopic versus mini-cholecystectomy, do not appear to provide definite evidence that the laparoscopic procedure is preferable. Furthermore, no data on long-term outcomes beyond 1 year seem to have been collected.

Evaluation of the economic aspects, of great importance in a resource-limited system, is posing equal, if not greater, problems than evaluations of clinical effects. The relative costs of laparoscopic and mini-cholecystectomy seem to vary depending on the site. The impact on costs of the changing pattern of demand and treatment have not been incorporated into financial calculations and, indeed, may not be entirely calculable. In addition, if economic analysis is an integral component of an RCT of a surgical procedure but the clinical questions are answered before the trial ends, it may not be possible to continue with the trial just to answer the economic questions. Hence, the evaluation of laparoscopic cholecystectomy overall has been less than satisfactory.

Timing

According to Bouchard and colleagues (1996), the timing and selection of a study are vital to its pertinence and feasibility, as "...the technology must be developed adequately to reflect a level of efficiency close to its optimal performance and needs to generate sufficient interest to justify the trial's time and expense while maintaining patient and investigator motivation". The time-span of some of the RCTs of laparoscopic versus mini- and open cholecystectomy are shown in *Figure 2* (page 23). As early as

1991, Neugebauer and colleagues claimed that the opportunity to conduct an RCT of laparoscopic versus open cholecystectomy had already passed, because clinical audit had by then shown the obvious benefits of the laparoscopic procedure, thus placing ethical constraints on the conduct of a controlled trial. Yet *Figure 2* shows that RCTs of both procedures, although admittedly small in terms of patient numbers, were still being initiated in 1991, suggesting that the window of opportunity for initiating such trials had not yet closed.

In the published reports of RCTs of laparoscopic cholecystectomy there is generally no indication of why a trial was undertaken at a particular time. An exception to this is Trondsen and colleagues (1993), who stated that, on the introduction of the procedure to their hospital in 1990, they took the opportunity of carrying out a prospective randomised study to compare laparoscopic with open cholecystectomy. Placing less emphasis on the importance of the learning curve prior to initiation of an RCT may have circumvented the ethical concerns which would have arisen from favourable clinical audit during this learning period. Starting the trial when the new technique was first introduced may also have meant that there was more genuine uncertainty as to the relative merits of the alternative treatments on offer.

In 1992, Barkun and colleagues stated that the introduction of alternatives to conventional cholecystectomy in their hospitals gave them a chance to carry out an RCT of laparoscopic versus mini-cholecystectomy. Surgeons participating in this trial, however, had already undertaken at least 30 laparoscopic procedures before the trial started. Here it seems that surgeons were able to experience the learning curve effect prior to an RCT without the trial being ethically compromised. McGinn and colleagues (1995) considered, when reporting their RCT of laparoscopic versus mini-cholecystectomy, that new techniques and operations, like new drugs, should continue to be evaluated by proper RCTs.

If it was possible to undertake small RCTs comparing laparoscopic with open cholecystectomy, then why not large ones? Perhaps one answer may lie in the results of the survey undertaken by McMahon and colleagues (1992), which showed that although there was wide support for such an RCT, those respondents with more laparoscopic experience were less convinced of the need for a trial. It seems that, in a situation where there was no centralised control of diffusion of a new surgical procedure, no consensus that promising

new techniques should not be allowed to diffuse until properly evaluated and no special training requirements for MAS techniques, surgeons could adopt laparoscopic cholecystectomy, gain experience with the technique, be impressed with the short-term outcomes and then conclude that an RCT was no longer necessary or appropriate. Meanwhile the procedure continued to spread into routine practice. In 1992, the Department of Health stated that if some technology has a large impact, observational data may suffice to demonstrate this effect clearly and carefully controlled trials are unnecessary. Perhaps surgeons with some experience of laparoscopic cholecystectomy came to believe that the procedure fell into such a category.

Surgeons' beliefs of how any new surgical innovation should be evaluated would also have played a part in the way in which laparoscopic cholecystectomy was assessed. A range of views exists, including, among others, randomisation of the first patient, randomisation following the learning curve, monitoring new techniques in selected centres, and a programmed evaluation containing an RCT as one of several elements.

Another factor which may have played a part in influencing when and what type of evaluation was undertaken was the existence of another alternative, mini-cholecystectomy. As a result, a number of RCTs were undertaken comparing laparoscopic with mini- but not open cholecystectomy.

The problem of fast-changing technology would have impacted on any decision about the timing of assessment. According to the NHS Executive (1996), a health technology under assessment should be sufficiently stable for it to be compared with competing health technologies and/or no intervention. Perhaps in the very early days of laparoscopic cholecystectomy it was considered that if an RCT was undertaken, developments in the procedure while the trial was in progress might compromise the validity of the results. However, once a technique has been allowed to diffuse into routine clinical practice, it becomes more difficult to conduct an adequate evaluation, and much more difficult to withdraw the technique – whatever the results of the trial.

Messages transmitted by the popular media are also powerful factors that can influence both when a technology is assessed and, at least indirectly, how it is assessed. There is general agreement in the literature that media reports of the apparent

benefits led to patients becoming aware of the procedure and asking clinicians for laparoscopic rather than open cholecystectomy. Patients may also have been persuaded not to take part in RCTs for fear of being randomised to the standard procedure. In addition, patient demand would have had an impact on the rate of diffusion, which would have implications for the timing of evaluation. However, later media reports of complications (see, for example, *The Times,* 8 December 1992) may have persuaded patients not to take part in any planned RCT for fear of being randomised to the laparoscopic procedure, thus influencing the way in which the technique could be evaluated. Such reports would also have turned the media spotlight on the (then) lack of specialised training requirements for surgeons who wished to undertake the procedure and, indeed, on the fact that an adequate evaluation of laparoscopic cholecystectomy was required. Thus the popular media may have had different effects at different times.

In the light of the papers reviewed, it is clear that systematic evaluation, in the form of an RCT, should have started from the earliest stages. In the case of such a radically new procedure, identification of its 'newness' should not have been problematic. Interpretation of the results of early evaluations, however, would have had to allow for the learning curve effect, and results of early trials would not necessarily correspond to results of trials undertaken by surgeons who had undergone well-designed training programmes before starting to use the procedure. In the light of this and the other problems affecting the quality of RCTs, a more tightly prescribed approach to reporting observational studies would improve the standard of information presented (while still acknowledging their limitations). Such a scheme – the CONSORT statement (Begg *et al,* 1996) – has recently been proposed in an attempt to improve the quality of reporting of RCTs.

The diffusion of laparoscopic cholecystectomy was not preceded by adequate evaluation of the technique. Many of the assessments in the literature are uncontrolled descriptive studies. There has been no large-scale RCT of laparoscopic versus open cholecystectomy. A few small RCTs have been undertaken and there have been a similar number comparing laparoscopic with mini-cholecystectomy. Observational studies and RCTs of laparoscopic versus open cholecystectomy generally reported a longer operation time, shorter hospital stay, less post-operative pain, faster return to normal activity and a much smaller scar. The most serious complication was an increased likelihood of injury to the bile duct, ranging from 0% to 4% in the

RCTs and from 0.6% to 1.8% in cohort studies and case series (Black, 1996). Early RCTs of laparoscopic versus mini-cholecystectomy found in favour of the former, although later RCTs concluded that it offered no clear advantage.

Few of the RCTs included any economic analysis. Those economic studies that have been undertaken suggested that laparoscopic is less costly than open but more expensive than mini-cholecystectomy. A complication is that costs seem to vary depending on the site and do not take into account the changing pattern of demand since the introduction of the laparoscopic procedure.

The timing of the evaluation of a new health technology is critical. There is a window of opportunity when it is possible to conduct an RCT but once the technology has diffused this becomes much more difficult to undertake. In retrospect, systematic evaluation of laparoscopic cholecystectomy should have started from the earliest stages of its introduction, but this did not happen. The reasons for this failure included the fact that there was no consensus to withhold new techniques until properly evaluated, neither was there centralised control to prevent the diffusion of unevaluated techniques. Favourable clinical audit during the learning curve period resulted in ethical objections to conducting an RCT, while the exact role of the RCT itself in the evaluation process was debated. In addition, media reports fuelled initial demand for the new treatment. Ironically, these also eventually highlighted the complications that resulted from surgeons adopting the technique without adequate training.

Chapter 5

Systematic review of diffusion and evaluation of CVS, focusing on factors influencing timing

Summary

The purpose of this systematic review is to investigate when and how evaluations of the prenatal diagnostic technique of CVS have been undertaken and what factors influenced the timing of those evaluations.

CVS and amniocentesis are the most common methods of prenatal diagnosis of chromosomal abnormalities. In CVS, cells are retrieved from the developing placenta for chromosomal analysis; this is usually undertaken, either transabdominally or transcervically, at about 9 weeks gestation in the first trimester of pregnancy, with a diagnosis available about 2 weeks later. In amniocentesis, transabdominal samples of foetal cells from the amniotic fluid surrounding the foetus are taken, usually at about 16 weeks gestation in the second trimester of pregnancy, with a diagnosis available about 3 weeks later.

Three major RCTs have been undertaken comparing first-trimester CVS with second-trimester amniocentesis, the aim being to measure pregnancy outcome, antenatal complications and diagnostic accuracy. The Canadian collaborative trial (1984–88) reported an excess foetal loss rate of 0.6% in the CVS group compared with the amniocentesis group but concluded that the total loss rates did not differ significantly between the two groups. The Danish RCT (1985–90) reported excess foetal loss rates for CVS compared with amniocentesis of –0.7% for transabdominal CVS and 4.6% for transcervical CVS; consequently, it was concluded that transabdominal CVS represented the most attractive option. The MRC European trial (1985–89) found that CVS reduced the chances of a successful pregnancy by 4.6% compared with amniocentesis. In a large US study (1985–86) undertaken by Rhoads and colleagues (1989), randomisation was attempted but had to be abandoned after only four women were randomised. Excess foetal loss rates for CVS were 0.8%. None of these trials incorporated any economic analysis and little was found in the literature on the costs of CVS.

A systematic review of CVS compared with amniocentesis for prenatal diagnosis by Alfirevic and colleagues (1995a) described these three major RCTs as being of generally good quality. Randomisation was organised centrally for all three trials, analysis on all randomised women was available for most principal measures of outcome, and the outcome of pregnancy was reported for all women in the Canadian Trial, 99% of women in the MRC European Trial and 93% in the Danish Trial. The review concluded that the results were consistent with the suggestion that second-trimester amniocentesis was safer than CVS. and therefore the benefits of earlier diagnosis by CVS had to be set against its greater risks. If earlier diagnosis was required, then transabdominal CVS was seen as preferable to the transcervical approach.

The question of a causal link between CVS and limb defects remained unresolved although a recent systematic review supports a causal relationship (Brown *et al,* 1996). Firth and colleagues (1991) were the first to report the possibility of such a link. None of the three major RCTs identified such a link, and a number of other studies produced conflicting results.

Attempts to start randomised trials early were successful in Canada, the UK and Denmark, to the extent that CVS was limited to the randomisation trials in Canada and some UK centres. In contrast, an attempt to mount a large randomised trial in the USA, at about the same time, failed. This reflects differences in physician and patient perceptions about the need for a randomised trial, the availability of CVS outside the trial, and the scientific and media exposure of the new technology.

Background

Prenatal diagnosis techniques can identify foetuses with a variety of genetic and non-genetic conditions. When these conditions are serious or potentially disabling, this knowledge allows parents to consider the option of terminating the pregnancy. Prenatal diagnosis

can be classified into three phases (Chitty & Bobrow, 1994):

(i) the identification of high-risk pregnancies that would benefit from prenatal diagnosis
(ii) the offering of the appropriate obstetric test, such as CVS or amniocentesis
(iii) laboratory testing, which may involve biochemical, cytogenetic or DNA analysis.

The first major concern with all prenatal diagnostic techniques is their accuracy. Weatherall (1991) pointed out a number of errors associated with prenatal diagnosis using foetal chromosomal DNA analysis. In CVS, contamination of foetal tissue with maternal tissue may occur, which may lead to mis-diagnosis. A number of technical problems in the analysis of DNA in the laboratory may give rise to errors, such as difficulty with DNA digestion and plasmid contamination (Weatherall, 1991). False-negative results occur if a genetic test fails to detect the specific form of gene when it is actually present, while false-positive results indicate wrongly that the specific form of gene has been detected. False-negative and false-positive results occur with different frequencies for different tests.

Chorionic villi consist of two types of cells: cytotrophoblastic cells, which are dividing rapidly and used for direct analysis, and mesenchymal core cells, which are used to initiate cell cultures. Whereas cytotrophoblasts can be directly analysed within 72 hours, the analysis of mesenchymal core cells usually takes 7–10 days (Shulman & Elias, 1993). Reliance on cytotrophoblastic cells alone is associated with false-positive results whereas testing both types of cell is costly and delays the results.

Chitty and Bobrow (1994) argued that while the accuracy of prenatal diagnosis depended primarily on the laboratory test, safety and sampling failure were dependent mainly on the obstetric procedure. The earlier in the pregnancy the procedure was carried out, the sooner the prospective parents could be informed of the results. Generally speaking, however, the earlier the test was undertaken, the more difficult it was to obtain an adequate sample for diagnosis and the more hazardous it was for the foetus, which would then be at a stage of rapid organ development. Even small differences in the safety or diagnostic accuracy of methods of testing would, potentially, have an important bearing on the choice of method for prenatal diagnosis (Chitty & Bobrow, 1994).

The second major concern is safety. The methods used to acquire material for prenatal diagnosis are all invasive. The vast majority of foetuses tested are, in fact, normal and do not stand to gain directly from prenatal diagnosis. The benefits of identifying abnormal foetuses have, therefore, to be set against any adverse effects of the procedure on normal foetuses. In this situation, even very small risks may be important.

The most common methods of prenatal diagnosis are amniocentesis and CVS. Amniocentesis is the transabdominal sampling of foetal cells from the amniotic fluid which surrounds the foetus and is usually performed at about 16 weeks gestation in the second trimester of pregnancy, with a diagnosis available about 3 weeks later. In CVS, cells are retrieved from the developing placenta for chromosomal analysis and it is usually performed after about 9 weeks gestation in the first trimester of pregnancy, with a diagnosis available about 2 weeks later. CVS can be undertaken either transabdominally or transcervically.

Development and diffusion

Schemmer and Johnson (1993) reported that Mohr introduced the concept of chorionic biopsy for foetal diagnosis in 1968. However, a combination of high complication and procedure failure rates, along with the rapid acceptance of amniocentesis effectively postponed any major European research in this area for approximately 10 years (Schemmer & Johnson, 1993). In 1970, a group of Chinese investigators also began to investigate the use of aspiration of villus material in an attempt to establish a safe and easy method for first-trimester sex determination, and reported on their series of 100 patients in 1975, correctly predicting the foetal sex in 94% of cases (Blakemore, 1988). Hod (1994) reported that the first CVS, in pregnancies that were allowed to go to term, was performed in China in 1983.

According to Modell (1985), the first experimental approach to CVS, in Scandinavia in the 1970s, was abandoned because of a high rate of complications, partly because obstetric ultrasound had not, at that time, been sufficiently developed. It was the addition of ultrasound guidance and advances in DNA technology which led to renewed interest in CVS in the early 1980s (Schemmer & Johnson, 1993). Visualisation of the placenta by ultrasound was found to improve the success rate of chorionic biopsy and soon afterwards investigators established that first-trimester villi could be used for DNA analysis in pregnancies at high risk for haemoglobinopathies and biochemical disorders (Schemmer & Johnson, 1993).

The reliability of transcervical aspiration at the time was hampered by the large diameter of the endoscopic instruments then in use and the inability to place the metal sampling cannula at an appropriate site. A smaller sampling device was developed, a soft 1.5 mm polyethylene catheter threaded with a malleable obturator which, with minor modification, remains the most frequently used method for transcervical CVS (Schemmer & Johnson, 1993). Simoni and colleagues (1983) then showed that villous material could be used for cytogenetic and biochemical diagnosis without the need for tissue culture (Schemmer & Johnson, 1993).

According to Froster and Jackson (1996), CVS was established as a method for prenatal diagnosis in 1982. It has been used clinically since early 1984 (Jackson *et al,* 1992) and the use of first-trimester transabdominal CVS was introduced by Smidt-Jensen and co-workers in 1984 (Schemmer & Johnson, 1993).

A WHO-sponsored International CVS Registry was opened in 1983 and, by the end of 1984, participating genetics centres around the world had reported summaries of over 3000 diagnostic cases (Modell, 1985). In the decade to 1995, CVS was used on more than 150,000 women worldwide for prenatal detection of genetic abnormalities (Olney *et al,* 1995).

Objectives

The objectives of this review were to identify when and how evaluations of CVS had been undertaken and to clarify the factors that influenced their timing.

The structure of the review is as follows. Following a brief section on search strategy and methods, the development and diffusion of CVS is discussed, followed by a description of the potential advantages and disadvantages of the technique. Major trials comparing first-trimester CVS with second-trimester amniocentesis are examined, as well as a systematic review of the RCTs. Transabdominal versus transcervical CVS is then reviewed, followed by CVS versus early amniocentesis. Economic evaluation, patient and physician preferences, and the issue of timing of evaluations follow, ending with a discussion of the findings.

Search strategy

Searches were carried out primarily on the databases Medline, Embase, and the Cochrane

Database of Systematic Reviews. These were supplemented with articles identified by other means, including contact with experts in the field, hand-searching of reference lists of key articles obtained and monitoring the contents of a small number of medical journals in 1996 (see Appendix 1 for a more detailed description of the objectives, criteria for inclusion, and methods).

Studies included

Potential advantages of CVS

The major advantage of first-trimester CVS over second-trimester amniocentesis is that it provides an earlier diagnosis (Chitty & Bobrow, 1994; Hod *et al,* 1994; Brambati, 1995). While CVS can be undertaken at about 9 weeks gestation, with the results available approximately 2 weeks later (Bryce *et al,* 1989), amniocentesis is usually performed between weeks 16 and 18, in the second trimester of pregnancy, with the results often not available until 2–4 weeks later (McGovern *et al,* 1986).

By detecting foetal abnormalities at an early stage CVS can potentially reduce maternal anxiety. Spencer and Cox (1987), in a study of 74 women who took part in the Canadian trial, noted that women who received CVS underwent a significant reduction in anxiety after prenatal testing. The anxiety level of women who had undergone amniocentesis, however, remained relatively high until they received the test results at about 19–20 weeks. Spencer and Cox (1988) argued that anxiety during pregnancy could be associated with subsequent obstetric complications, including somatic complaints and labour and delivery problems.

Women who underwent CVS also reported less physical discomfort than women having amniocentesis (Spencer & Cox, 1987).

If the results of CVS revealed a serious or disabling foetal abnormality, the woman could choose to undergo a safer, less traumatic, first-trimester therapeutic abortion (Heckerling *et al,* 1994). Couples may also feel that it is morally more acceptable to terminate a pregnancy in the first, rather than in the second, trimester (Modell, 1985). According to Spencer and Cox (1988), during the second trimester women normally experienced an increasing foetal presence and a growing attachment to the foetus. A second-trimester abortion, therefore, carried increased psychological trauma, as well as increased maternal morbidity and mortality (McGovern *et al,* 1986).

For the purposes of DNA and biochemical analysis, CVS afforded immediate access to significant quantities of tissue without the need for cell culture as was required following amniocentesis (Chitty & Bobrow, 1994). Also, rapid analytical techniques had significantly reduced the waiting time between sampling and diagnosis, and progress in recombinant DNA technology and human gene mapping had led to an increase in the range of conditions that could be detected (Brambati, 1995).

Potential disadvantages of CVS

Initially, the principal concern was the risk of provoking miscarriage in normal pregnancies. A variety of other short-term risks were identified from the case studies or case reports, including concerns about serious infection. At this time, there was no suggestion that CVS might cause congenital limb reduction deformities; such concerns were prompted later, in 1991, by a series of five cases among the CVS procedures performed at a single centre, Oxford (Firth *et al,* 1991). This is discussed further below.

RCTs of CVS

The literature search identified three major RCTs comparing first-trimester CVS with second-trimester amniocentesis.

Canadian Collaborative Trial

The Canadian Collaborative CVS–Amniocentesis Clinical Trial Group (Lippman *et al,* 1992) reported that 2787 women aged 35 years or older at the expected date of delivery, were randomised either to transcervical CVS at 9–12 weeks gestation or to amniocentesis at 15–17 weeks. The trial, covering 1984–89, was designed to compare foetal loss rates, including induced abortion, and 11 centres participated. The accuracy of both procedures was also examined. Participating obstetricians were required to have carried out at least 30 procedures and to have been successful in obtaining 10 mg or more of chorionic villi in 23 of 25 consecutive cases, thus indicating the importance attached to the experience and skill of the practitioner.

The results of the Canadian trial showed an excess foetal loss rate of 0.5% in the CVS group compared with the amniocentesis group. The potential difference in loss rates between CVS and amniocentesis was believed to be not more than 2.4% for women of 35 years of age and over with a viable foetus at the time of the procedure. It was suggested that the results might reassure women on the safety of first-trimester CVS. The Canadian trial group concluded that the total loss rates were not statistically significantly different between the groups, irrespective of whether data on all women randomised or only on those women eligible were analysed.

Lippman and colleagues (1992) noted that although there were very few false-negative results with either CVS or amniocentesis (3/24 confirmed true-positives and 0/20 confirmed true-positives, respectively), false-positive results occurred more often with CVS than with amniocentesis (19 positive results from 839 confirmed true-negatives and two from 947 confirmed true-negatives, respectively). Lippman and colleagues (1992) stated that the less accurate CVS diagnoses were caused primarily by 'confined mosaicism' – chromosome abnormalities confined to the placental villi and not found in the foetus proper. There might also be a higher frequency of maternal contamination in the CVS specimens. They concluded that although the results suggested that prenatal cytogenetic diagnosis by CVS and amniocentesis were both very accurate overall (97.5% and 99.8%, respectively), the difference was statistically and clinically significant.

Medical Research Council European Trial

The MRC Working Party on the Evaluation of Chorion Villus Sampling (1991) recruited 3248 women seeking prenatal diagnosis to an international, multicentre, randomised trial of CVS versus amniocentesis which ran from 1985 to 1989. The aim of the trial was the reliable identification of any differences in the safety or diagnostic accuracy of methods of testing between first-trimester CVS and second-trimester amniocentesis, when used in everyday clinical practice.

The trial was coordinated by the National Perinatal Epidemiology Unit in Oxford and the MRC Human Genetics Unit in Edinburgh. In all, 31 centres took part – from the UK (21), Italy (4), The Netherlands (2), Finland (1), Denmark (1), Switzerland (1), and Germany (1). A centre was eligible to take part if each participating obstetrician had carried out 30 or more CVS procedures, as it was considered that the safety of the technique might be positively affected by the skill of the clinician improving with experience. At some centres, CVS was made available only in the context of the trial, a policy recommended by the Working Party on the grounds that CVS was a new technique with unknown benefits and

disadvantages, and therefore should not be made routinely available until it had been adequately evaluated. The results are detailed below.

1. CVS was carried out in whatever way was deemed suitable by the obstetrician (72% by the transcervical and 28% by the transabdominal approach).
2. The results suggested that CVS in the first trimester reduced the chances of a successful pregnancy by 4.6% compared with amniocentesis in the second trimester.
3. The trial provided no clear evidence that the choice of route in CVS affected the risk of spontaneous pregnancy loss. The MRC Working Party qualified this by acknowledging that only indirect comparisons were made between transcervical and transabdominal CVS in the study and the most reliable information would probably come from RCTs comparing the two methods directly.
4. The MRC Working Party reported that three terminated pregnancies were false-positives (one tested by CVS and two by amniocentesis), and that two other mosaic cases diagnosed by CVS may have been false-positives. In addition, there was one CVS false-negative result.
5. There were 38 congenital malformations in liveborn or stillborn infants in the CVS group and 41 in the amniocentesis group. The numbers with each malformation were small and there were no clear differences between the groups in patterns of malformation. Although there were two cases of limb reduction deformities in the CVS group, both were in the case series reported from Oxford. The MRC Working Party therefore argued that the European trial provided no new information that either confirmed or refuted any suggested link between early CVS and limb abnormalities.

Danish trial
Smidt-Jensen and colleagues (1992) reported a Danish RCT comparing transcervical CVS, transabdominal CVS and second-trimester amniocentesis; this took place from August 1985 to November 1990. The trial was conducted at two centres, one in Copenhagen and the other a provincial centre in Sonderborg. Most patients (85%) were recruited in Copenhagen and, for the majority of those who were allocated to CVS, the operation was performed by a single experienced operator. Women had no access to CVS at the two centres unless they agreed to join the trial. However, because other hospitals in Copenhagen offered CVS, the conduct of the study was made more difficult than it would otherwise have been.

According to Smidt-Jensen and colleagues (1992), nine women allocated to the amniocentesis arm elected instead to visit another hospital for CVS.

Participating obstetricians were required to have carried out at least 20 successful samplings of chorionic villi and amniotic fluid before the Danish trial, a lower number than in the Canadian and MRC European trials (30). A total of 4758 women were referred to the trial, of whom 4214 were regarded as being at low genetic risk; the total number randomised was 4199, of whom 3706 were at low genetic risk. Of the patients who completed the study, 1429 had transcervical CVS (1010 at low genetic risk), 1453 had transabdominal CVS (1027 at low genetic risk), and 1115 had amniocentesis (1042 at low genetic risk).

Total foetal loss rates for all low genetic risk women were reported for the Danish trial as 11.6% for transcervical CVS, 6.3% for transabdominal CVS and 7% for amniocentesis. The most important difference between the transabdominal and transcervical CVS groups was considered to be the proportion of post-procedure unintended losses, the two groups having otherwise comparable foetal losses.

According to Smidt-Jensen and colleagues (1992), the risk of foetal loss was similar after transabdominal CVS and amniocentesis but, because losses after amniocentesis came at a later stage, they were more distressing. In May 1990 recruitment to the amniocentesis group stopped because of increasing difficulty in assigning women to that procedure and also because of the apparently small difference in foetal loss between transabdominal CVS and amniocentesis. Randomisation of the two CVS approaches continued until November 1990, when it was decided to abandon transcervical CVS at both trial centres on the grounds that this procedure posed a greater risk to the foetus.

Smidt-Jensen and colleagues (1992) reported that, in women at low genetic risk, the proportion for whom a cytogenetic diagnosis was obtained at the first attempt was 96% for transcervical CVS, 98.1% for transabdominal CVS and 99.7% for amniocentesis.

The distribution of congenital abnormalities between the groups was similar and there were no differences in the distributions of stillbirths or neonatal deaths.

Smidt-Jensen and colleagues (1992) argued that transabdominal CVS allowed better access to the

placental site than transcervical sampling, had the potential that more villi could be aspirated when needed, was an easier skill to acquire, was more dignified, required less time and was more acceptable to women than the transcervical approach. They concluded that the results of their comparisons of the three procedures showed transabdominal CVS to be the most attractive choice because it offered an early and acceptable resolution of anxiety with little foetal and maternal risk.

Observational study of CVS

In a large, seven-centre study in the USA carried out during 1985 and 1986, Rhoads and colleagues (1989) attempted to compare the safety and efficacy of CVS in 2278 women with that of amniocentesis in 671 women. Both groups were recruited in the first trimester of pregnancy and had viable pregnancies verified by ultrasound examination. The collaborative study was initiated to address issues such as foetal loss rates after CVS and to determine more exactly the rates of complications involved.

The centres collaborating in the study were selected because of their members' interest in CVS. Although the clinicians' experience with the procedure varied considerably at the outset, each clinician had by then completed at least ten CVS procedures. In addition, each centre had completed at least 25 procedures and was routinely offering both CVS and amniocentesis to women who presented with indications for prenatal diagnosis (Rhoads *et al,* 1989).

Although at the outset the study design was intended to be that of a randomised trial, this approach had to be abandoned. Rhoads and colleagues (1989) reported that despite a vigorous effort in the first 4 months of the study, only four women agreed to be randomly assigned to either CVS or amniocentesis. This led to transcervical CVS becoming the standard prenatal diagnosis method of the trial, and only those women (23%) who definitely wanted amniocentesis were assigned to the control group. Rhoads and colleagues (1989) qualified the results of their trial by conceding that the absence of randomisation could have produced falsely reassuring results in the CVS group if the loss rate in the amniocentesis group had been unusually high. Even in the absence of any information about potential side-effects, women apparently insisted on choosing the method that reduced anxiety and was more comfortable. (A similar situation existed at a comparable stage in the development of laparoscopic

cholecystectomy, when patients favoured the laparoscopic procedure with its perceived benefits of a smaller scar, less postoperative pain, and a faster return to normal activity.)

Rhoads and colleagues (1989) argued that the safety concerns of women who clearly wanted prenatal diagnosis were related to the perceived extent of any excess risk to the foetus associated with CVS as opposed to amniocentesis. They reported that their best estimate of the excess risk was 0.8% which, although not statistically significant, nevertheless did suggest a degree of excess risk with CVS. As they had found that foetal loss rates reported to the CVS Registry had varied quite widely, they also recommended that transcervical CVS should be attempted only at medical centres that planned to undertake substantial numbers of procedures, suggesting that they considered that the level of experience affected the success of the outcome. They noted that the 97.8% success rate in making CVS cytogenetic diagnoses did not quite match the 99.7% success rate for amniocentesis but commented that the difference was small and would probably be acceptable to many patients.

Cochrane systematic review of CVS compared with amniocentesis for prenatal diagnosis

A search of the Cochrane Database of Systematic Reviews identified a systematic review of CVS compared with amniocentesis for prenatal diagnosis (Alfirevic *et al,* 1995a) based on the three major RCTs referred to above.

The objectives of the review were to compare the safety and accuracy of first-trimester CVS with second-trimester amniocentesis by testing the following three hypotheses:

(i)   that CVS, irrespective of the route and instrument used, was as safe and accurate in obtaining correct prenatal diagnosis as second-trimester amniocentesis

(ii) that CVS by the transabdominal route was as safe and accurate in obtaining correct prenatal diagnosis as second-trimester amniocentesis

(iii) that CVS by the transcervical route was as safe and accurate in obtaining correct prenatal diagnosis as second-trimester amniocentesis.

Outcome measures included foetal and pregnancy outcomes, antenatal complications and diagnostic accuracy.

The reviewers stated that the outcome of pregnancy was reported for all women in the

Canadian trial, 99% of women in the MRC European trial, and 93% in the Danish trial. Where data were available for all three trials the results were consistent, particularly for data on pregnancy outcome. CVS appeared to be more technically demanding for both obstetricians and cytogeneticists. Complications were uncommon after both procedures and there were no reports that these were ever life-threatening. Importantly, however, pregnancy loss was more common after allocation to CVS. The suggestive increase in stillbirths and neonatal deaths following CVS, although not statistically significant, was nevertheless described as worrying.

Transcervical CVS yielded more abnormal karyotypes than amniocentesis, which resulted in a higher number of terminations of pregnancies in the transcervical CVS group. There were also significantly more pregnancy losses following transcervical CVS, with the increase in the number of spontaneous miscarriages reaching statistical significance. The same trend was observed in the numbers of perinatal deaths for the transcervical CVS group.

Data for a comparison of transabdominal CVS versus amniocentesis were available only from the Danish trial and for a small number of outcomes. There were significantly more abnormal karyotypes in the transabdominal CVS group which resulted in the trend towards more terminations of pregnancy in that group. There was no differential effect on the total pregnancy loss, spontaneous loss before viability, number of stillbirths and neonatal deaths, and number of congenital abnormalities (Alfirevic *et al*, 1995a).

The question of whether the clinicians were sufficiently skilled to undertake CVS satisfactorily remained controversial. In all three trials, operators were required to have performed CVS successfully at least 20 times in order to participate. The reviewers concluded that there was no clear evidence that performance improved over the course of the RCTs, and questioned whether even very skilled operators could improve the performance of CVS enough to abolish the difference in pregnancy loss between the two groups.

Alfirevic and colleagues (1995a) stated that none of the trials was designed to assess the diagnostic accuracy of prenatal testing adequately. This question, therefore, remained unanswered and the hypothesis that both CVS and amniocentesis were equally accurate remained untested. Nevertheless, the available data did suggest that accurate

diagnosis was more likely following amniocentesis. The authors emphasised that, although absolute numbers of false-positive and false-negative results were small, they can have such devastating effects that observed differences should not be ignored.

As far as the implications for practice were concerned, Alfirevic and colleagues (1995b) concluded that second-trimester amniocentesis was safer than CVS and therefore the benefits of earlier diagnosis by CVS had to be set against its greater risks. If earlier diagnosis was required, they argued that transabdominal was preferable to transcervical CVS. They pointed out the importance of prospective parents considering prenatal diagnosis being fully informed of the risks and benefits of alternative procedures before making a choice.

Alfirevic and colleagues (1995b) concluded that alternative methods for early prenatal diagnosis had to be sought in view of the high foetal loss rates and diagnostic inaccuracies of CVS. In addition, they recommended that any such new methods of prenatal diagnosis should be rigorously evaluated before a decision was made about their introduction into clinical practice. They suggested that in future trials, to the assess safety and accuracy of new methods, amniocentesis performed after 15 weeks should be considered as a control.

Transabdominal versus transcervical CVS

When first introduced into clinical practice, CVS was performed by transcervical catheter aspiration, a method which was adopted rapidly at many centres in Europe and North America (Jackson *et al*, 1992). It was argued that several clinical trials (USA multicentre trial; Canadian trial; Danish trial; MRC European trial) had found the transcervical approach to be both safe and efficacious for prenatal diagnosis. When a number of conditions were identified, however, for which the transcervical route appeared to be inappropriate, the transabdominal approach became increasingly adopted as an alternative (Brambati *et al*, 1991). The advantages claimed for the transabdominal approach included a lower risk of infection, ease of learning because of its similarity to transabdominal amniocentesis, and overall increased safety from sampling by needle rather than a plastic catheter (Jackson *et al*, 1992).

Brambati and colleagues (1991) undertook an RCT from March 1986 to July 1988 involving 1194 women randomised at 7–12 weeks' gestation. The purpose of the trial was to evaluate the relative advantages and disadvantages of transabdominal

and transcervical CVS in terms of foetal risks and efficacy. Despite the fact that they had previously used the transcervical route exclusively, these authors stated that pressure on the clinicians, arising from the widespread shift towards transabdominal CVS, eventually led to discontinuation of recruitment. Brambati and colleagues argued that the trial appeared to provide reassuring evidence in respect of the risks associated with both sampling techniques and concluded that transabdominal and transcervical CVS appeared to be equally effective.

Jackson and colleagues (1992) carried out a randomised trial at eight centres in the USA from April 1987 to September 1989. The aim was to confirm or disprove the presumed advantages of the transabdominal approach. The trial involved 3999 women with singleton pregnancies, in whom the risk of a genetically abnormal foetus was increased; 2010 were assigned to undergo transcervical sampling and 1989 transabdominal sampling. According to Jackson and colleagues, the two groups were similar and representative of the women in the USA who were then seeking prenatal diagnosis and they concluded that there were only minor differences between the two procedures.

Despite the findings of Jackson and colleagues (1992) and Brambati and colleagues (1991), Chitty and Bobrow (1994) argued that comparisons of the transabdominal and transcervical routes had in fact indicated conflicting results. They referred to the Danish RCT, where the rates of unintentional loss after CVS were 7.7% for the transcervical approach and 3.7% for the transabdominal approach (Smidt-Jensen *et al,* 1992). In addition, there was a significantly greater number of sampling failures in the Danish transcervical group. The systematic review of CVS compared with amniocentesis for prenatal diagnosis, undertaken by Alfirevic and colleagues (1995a) and included in the Cochrane Database of Systematic Reviews, also supported the results of the Danish trial which showed that transabdominal CVS appeared to be safer than transcervical CVS.

Putative risk of limb reduction deformity following CVS

In early 1991 a high rate of limb defects among children exposed to CVS in early pregnancy was reported from a cluster observed by Firth and colleagues (1991); Burton and colleagues (1992) subsequently reported a second cluster. The initial observations by Firth and colleagues suggested that the highest risk of limb malformations was associated with CVS being undertaken at a very early stage in gestation. Mastroiacovo and colleagues

(1992), in a case–control study (January 1988 to December 1991) using data from the Italian Multicentre Birth Defect Registry, reported that the risk of limb defects was increased following CVS, and that this increased risk applied especially to CVS performed before 70 days gestation; 131 hospitals participated and 423,087 births were studied.

Other studies, however, have found no connection between CVS and limb defects. Kaplan and colleagues (1990) evaluated 189 infants whose mothers had either CVS or amniocentesis as part of the Canadian trial. They concluded that there was no association between CVS and limb abnormalities. The analysis by Froster and Jackson (1996), based on 138,996 CVS outcomes, concluded that the overall incidence of limb defects in the CVS cohort did not differ from that in the general population. They also reported no correlation between gestational age at CVS and severity of defects. Evans and Hamerton (1996) criticised this study as being flawed, in that the two populations (CVS registry group and British Columbian cases) were not strictly comparable, data were collected over different periods and the quality of data was potentially variable. Brambati (1995), in an analysis of limb reduction defects among more than 130,000 cases reported to the WHO CVS registry, was unable to find any relationship between sampling and foetal malformations, including limb reduction defects. Thus, these studies failed, collectively, to give a clear indication of whether there might be an association under some conditions but not others.

In 1992, however, a meeting organised by the WHO Regional Office for Europe had recommended that CVS should be performed between 9 and 12 weeks of gestation (Olney *et al,* 1995). The Report of the National Institute of Child Health and Human Development Workshop on CVS and Limb and Other Defects (20 October 1992) recommended that further studies were needed to determine whether there was a causal link between CVS and limb defects, and whether the timing of the procedure was an important risk factor (Olney *et al,* 1995). Orrell and Lilford (1990) argued that it would be difficult to confirm rare potential side-effects by randomised studies alone, as only large increases in rare side-effects could be detected in this way. However, Evans and Hamerton (1996) argued that an ideal study would take the form of a large international RCT but added that it was now unlikely that this would be attempted.

Olney and colleagues (1995) undertook a multistate case–control study in the USA of

421,489 births from 1988 to 1992 in order to assess and quantify the risk of specific limb deficiencies associated with CVS. The study was instigated because a number of infants had been reported with transverse limb deficiencies after their mothers had undergone CVS but it had been unclear whether the procedure itself had caused the defects. The implication was that if they were not caused by CVS, then they were either occurring naturally or had some other cause.

The 'cases' were 131 infants with non-syndromic limb deficiencies ascertained from seven population-based birth-defect surveillance programmes, born between 1988 and 1992 to mothers aged 34 years or older. The controls were 131 infants with other birth defects. Exposure to CVS in the seven states was associated with a six-fold increase in the risk of transverse digital deficiency. The association was strongest when the procedure had been performed before 10 weeks' gestation. However, five of seven digital deficiencies occurred after procedures at 10 weeks gestation or later (Olney *et al,* 1995). The authors concluded that there was an increased risk of transverse digital deficiencies after CVS but that, as limb reduction deformity was rare, it would have occurred relatively infrequently in centres that had carried out less than 10,000 procedures.

Olney and colleagues (1995) argued that the case–control design of their study was advantageous because of the rarity of the outcome they were investigating. They also claimed that the population-based nature of the study made the findings more generalisable than would have been the case with studies confined to specific centres. They qualified their conclusions, however, by stating that, given the observational nature of the study, it would be virtually impossible to rule out the impact of unmeasured confounding factors; for example, they stated that women who underwent CVS differed from those who did not with respect to sociodemographic, cultural, and life-style factors which could potentially act as confounding variables (Olney *et al,* 1995).

Brown and colleagues (1996) have recently completed a systematic review of the literature and, despite apparent inconsistencies between some studies, they concluded that there is good evidence of an association between CVS and congenital limb-reduction deformity.

CVS versus early amniocentesis
The MRC Working Party on Chorion Villus Sampling (1991) had stated that earlier

amniocentesis, undertaken at 10–14 weeks gestation, potentially represented an attractive alternative to CVS. However, as its diagnostic accuracy and safety were not known reliably, it was recommended that early amniocentesis should be subjected to the same rigorous evaluation as first-trimester CVS, with no introduction into clinical practice before the results of such an evaluation could be made available.

A search of the Cochrane Database of Systematic Reviews identified a systematic review of early amniocentesis versus CVS undertaken by Alfirevic and colleagues (1995b). The authors pointed out that one of the major disadvantages of 16-week amniocentesis was that a final result was normally not available until after 18 weeks gestation, and that such a long wait could be very distressing for the couples involved. A possible alternative method might be that of early amniocentesis, carried out at 9–14 weeks gestation. This technique was very similar to routine amniocentesis except that the available pool of amniotic fluid was smaller (making it technically more difficult) and less amniotic fluid was removed making laboratory testing more difficult (Alfirevic *et al,* 1995b). The objectives of the systematic review were to compare the safety and accuracy of early amniocentesis with CVS by testing the hypothesis that amniocentesis at 10–13 weeks gestation was as safe and accurate, in obtaining the correct prenatal diagnosis of foetal karyotypes, as transabdominal CVS performed during the same period.

The Cochrane review was based on the data from the King's College Hospital Trial (Byrne *et al,* 1991). Alfirevic and colleagues (1995b) also noted a trial undertaken in Leiden but not included in the review because it had not been fully reported and many details were unavailable. The King's College Trial indicated that pregnancy loss was more common after allocation to early amniocentesis, and that the increase in total pregnancy loss could be ascribed both to an increase in spontaneous loss before 24 weeks and to an increase in perinatal mortality after 24 weeks. It was thought, however, that early amniocentesis was probably technically less demanding than transabdominal CVS, so that the frequency of undertaking transabdominal CVS might have an effect on the levels of pregnancy loss.

Alfirevic and colleagues (1995b) argued that although data from the King's College Hospital trial suggested an important increase in pregnancy loss following early amniocentesis, it was not possible to determine the true size of the effect

as the data were from one trial with only 488 participants. They argued that the safety and accuracy of prenatal invasive procedures might depend on gestational age because of factors such as changes in the size of the placenta, the amount of amniotic fluid, the number of viable foetal cells and the size of the coelomic cavity before it is obliterated by the fusion of amnion and chorion. They argued that it was also generally accepted that the number of complications decreased with advancing gestational age.

These authors drew attention to a particular concern about early amniocentesis, namely a possible adverse effect on lung development. They pointed out that as amniocentesis at about 16 weeks was known to increase the risk of neonatal respiratory problems, earlier amniocentesis might have a more serious effect because a larger proportion of the total fluid was withdrawn. Nevertheless, according to Alfirevic and colleagues (1995b), the available data suggested that early amniocentesis did not carry a greater risk than CVS performed at the same gestational age.

Alfirevic and colleagues (1995b) argued that, given the continuing concern about the safety and diagnostic accuracy of early amniocentesis and CVS, amniocentesis at about 16 weeks should be considered the routine method for prenatal diagnosis. They recommended that earlier use should be limited to specific circumstances where earlier diagnosis is required or to participation in further RCTs.

Economic evaluation of CVS

The three major RCTs (Canadian, MRC European, Danish) and the large US study did not include any financial analysis of the costs and benefits of CVS. However, Mugford and colleagues (1985) pointed out that the economic aspects of CVS and amniocentesis would have to form part of the decision-making process. Incorporating an economic evaluation as part of an RCT, rather than bolting it on afterwards, would make the trial results more accurate, valid, and easier to analyse.

Heckerling and Verp (1994) noted that an Italian study (Marchese *et al,* 1986) had found that when the costs of the sampling procedure, laboratory testing and spontaneous and therapeutic abortions were considered, the expected costs of CVS were 22% lower than those of amniocentesis. However, assumptions concerning costs were not tested by sensitivity analysis, so the generalisability of the results remained uncertain. Also, because effectiveness was not examined in the Italian

study, comparisons of cost per abnormal birth averted and cost per quality-adjusted outcome could not be made for the two procedures.

Heckerling and Verp (1994) carried out an economic study of amniocentesis and CVS for prenatal genetic testing (in the USA). Decision analysis was used to compare outcomes, costs, and cost-effectiveness of the two tests for a model cohort of 100,000 pregnant women aged 35 years at their expected date of delivery. The model assumed that prenatal genetic testing would be performed solely for the indication of maternal age. Utilities were used for the quality adjustment of outcomes of prenatal testing, with a maximum utility of 1 assigned to the birth of a chromosomally normal child and a minimum utility of 0 to the birth of an abnormal child. Other outcomes, such as spontaneous abortion, therapeutic abortion and maternal morbidity were assigned utilities between 0 and 1.

Hospital and physician costs for prenatal services were obtained from University of Chicago Hospitals' charge data, with all charges in 1992 USA dollars. Direct medical costs were included, such as the costs of prenatal care, amniocentesis, CVS, spontaneous abortion, therapeutic abortion and delivery. Indirect costs, such as time lost from work for prenatal testing, were not included.

According to Heckerling and Verp (1994), based on costs per abnormal birth averted, at all maternal ages from 30 to 43 years, amniocentesis was more cost-effective than CVS; at ages 44 and 45 years, CVS was more cost-effective. However, if the anxiety reduction provided by first-trimester diagnosis was equivalent to a 0.2% risk of an abnormal child, then CVS was more cost-effective than amniocentesis at all maternal ages. Based on data from the 1988 US natality cohort, the policy of testing women aged 35 and older would cost $103,329 and $111,184 per abnormal birth averted for amniocentesis and CVS, respectively. If women aged 30 and over were tested, this would almost double the cost. The authors argued that for either prenatal test, targeting high-risk women for testing, and striving for utilisation rates of 50% or higher, appeared to be the most cost-effective policy.

In conclusion, Heckerling and Verp (1994) argued that amniocentesis was more cost-effective than CVS at most maternal ages, although the differences in cost per abnormal birth averted were relatively small. However, when testing was undertaken at older maternal ages, and intangible benefits such as a reduction in anxiety due to the availability of earlier diagnosis were

considered, then CVS could become the more cost-effective procedure.

Media reports

Material on CVS retrieved by the Promt search included articles in *Newsweek* and the *Financial Times.* The report in *Newsweek* (22 June 1992) drew attention to the fact that CVS had been implicated in birth defects in a number of studies, adding that some hospitals had put restrictions on offering CVS. It suggested that the procedure was best performed at facilities with considerable experience in CVS and concluded that alternative, less invasive measures for genetic sampling of foetal cells were under development but were years away from clinical trial status. The *Financial Times* (12 January 1996) forecast that, within 2 or 3 years, a single blood test could determine whether a foetus has a condition such as Down's Syndrome, multiple sclerosis or cystic fibrosis. The report pointed out that although CVS could be performed at an earlier stage than amniocentesis, it carried a higher rate of miscarriage and that, for both procedures, there was a chance of physical damage to the baby.

These articles may have discouraged some women from undergoing CVS who would otherwise have done so. Other reports appeared in more technical journals, detailing companies that were developing new techniques for prenatal testing.

Timing of evaluations

The reports of the three major RCTs (Canadian trial, MRC European trial, Danish trial), and of the large USA collaborative study in which randomisation had to be abandoned, provide general information on the reasons why the evaluations were carried out at a particular time. The Canadian Collaborative CVS–Amniocentesis Clinical Trial Group (1989) stated that their multicentre randomised trial was started in 1984 to ensure that CVS was not introduced in Canada before its risks and safety had been assessed. As has been previously noted, this trial was remarkable in that CVS was only available in Canada to women who participated in the study. Evaluation therefore started early, soon after first introduction of CVS into clinical practice, and diffusion was limited to trial centres until the trial was completed.

The MRC Working Party on the Evaluation of Chorion Villus Sampling (1991) initiated their European collaborative trial in 1985 because of the emergence of first-trimester CVS as an alternative to second-trimester amniocentesis; although CVS had the advantage of allowing earlier diagnosis its safety and diagnostic accuracy was not yet known.

Early discussions between groups considering local trials and a national research centre specialising in perinatal HTA led to the rapid establishment of an MRC Working Party, which supported the development of a trial protocol. Randomisation started sufficiently early for many participating centres to limit the availability of CVS to participants in the trial. Another key factor behind the success of this trial was the active involvement and support of the relevant consumer groups.

The Danish randomised study, which compared the safety, efficacy and accuracy of transabdominal CVS, transcervical CVS, and amniocentesis, began in August 1985. The trial was limited to a small number of operators with particular expertise in transabdominal CVS. Since CVS was a recent technical innovation, they wanted the methods to be critically tested before they were routinely used. The authors argued that medical ethics required thorough investigation of the risks associated with new procedures before they were introduced for routine use, while stating that figures for the learning periods would inevitably reduce the apparent efficacy of the innovation.

In the USA, attempts to mount a large RCT also started in 1985 but this approach had to abandoned as the vast majority of women were reported to be demanding CVS. In both the USA and Italy, where randomised comparison with second-trimester amniocentesis had been planned, investigators instead chose to concentrate efforts on randomised comparison of transcervical with transabdominal CVS, starting in 1987 and 1986, respectively.

Studies of congenital limb reduction deformation were prompted later by the report of Firth and colleagues (1991). Quite rightly (Chalmers, 1987), many investigators chose to use the case–control design which is especially suitable for studying rare outcomes. These used a retrospective design to compare cases and controls in respect of their exposure to CVS.

The timing of the major trials in relation to CVS references in Medline

The annual number of references to CVS identified by a search of the Medline database are shown in *Figure 4* (page 25), together with numbers of CVS conferences by year, identified by searching the Index of Scientific and Technical Proceedings and Medline. Against these are set the dates when major trials were carried out – the Canadian collaborative trial, the Danish trial, the MRC European trial, and the large USA trial in which randomisation was abandoned.

References to CVS first appeared in 1983, gradually increasing until they reached a plateau between 1989 and 1993, since when the annual numbers of references have been steadily declining. This decline suggests that CVS no longer arouses the same level of interest as it once did in the medical literature and that it has either become established in routine practice or, alternatively, is being used less. Note that the annual number of references to CVS peaked several years after it had diffused into routine practice.

Most of the data on CVS conferences was obtained from the Index of Scientific and Technical Proceedings. Although not an exhaustive list, it nevertheless provides a useful indicator as to the numbers of conferences taking place over this period, where papers on CVS would have been delivered, the technique discussed and information exchanged. *Figure 4* shows one conference in 1983, four in 1984, and six in 1985, with the numbers from 1986 to 1992 varying between three and five conferences annually.

Figure 4 also shows that the major randomised trials, and the large USA study, took place during the period when the number of publications on CVS was steadily increasing (with the exception of 1987). The Canadian trial began in 1984, the year after the first references to CVS appeared, with the other major trials starting in 1985. In the USA, an attempt to mount an RCT failed even though this was only 2 years after the first CVS reports.

Patient preferences

Lippman and colleagues (1985) reported a survey of women's attitudes to CVS carried out at McGill University, Montreal. The survey was intended to identify those aspects of the procedure that made it more (or less) acceptable than amniocentesis to women who were eligible for prenatal testing and to provide data to help estimate the numbers of women likely to participate in the planned Canadian controlled trial.

In the absence of precise estimates of CVS-associated risk at the time of the survey, almost equal proportions preferred amniocentesis and CVS. Whereas risk information was the most important factor to women preferring amniocentesis, the timing of the test or nature of the termination procedure was most important to those preferring CVS. The data suggested that the ultimate acceptability of the new procedure, by women aged over 35 years seeking prenatal diagnosis, would depend on the risk associated with it. When CVS was stipulated to be only as

'risky' as amniocentesis, most women (82%) expressed a preference for it. However, when the risk of foetal loss following CVS was stipulated to be 5% greater than that following amniocentesis, less than a quarter of the women (22%) continued to prefer the new technique. McGovern and colleagues (1986) carried out a questionnaire study of 520 women who had previously undergone amniocentesis. Respondents who chose amniocentesis did so because of the known low risk of spontaneous abortion, while for those who chose CVS the major criterion was that it was performed in the first trimester. Similar experiences were reported in a study by Bryce and colleagues (1989).

Lippman and colleagues (1985) argued that since the standard procedure for women seeking prenatal diagnosis was amniocentesis, only women preferring CVS would be likely to enrol in an RCT if it was not otherwise available. This was contradicted by the experience of the large USA collaborative study, where randomisation proved impossible because of patient reluctance to participate in a trial in which there was a 50% chance that they would be assigned to receive amniocentesis. Here too, given equivalent efficacy, cost and safety, both pregnant women and their physicians would have preferred first-trimester to second-trimester prenatal diagnosis of foetal abnormalities (Rhoads *et al,* 1989).

Hamerton (1989) stated that as a result of an agreement between centres delivering prenatal diagnosis in Canada, CVS was only obtainable in Canada through participation in the collaborative trial. Thus women allocated to amniocentesis and who found that unacceptable would need to travel to the USA to obtain CVS, which did happen in a few cases. In contrast, Muggah and colleagues (1987) stated that fear of a possible increased risk of loss associated with CVS was the overwhelming reason women gave for declining to participate in the trial. However, most women who entered the study accepted the rationale for randomisation but were often disappointed when assigned to the amniocentesis group.

Abramsky and Rodeck (1991) described a study in which 580 women were offered the chance to join the MRC European trial and were given four options: no test; definite amniocentesis; definite CVS; or randomisation between amniocentesis and CVS. They noted that CVS was more popular than amniocentesis, timing being the reason why women preferred it. In addition, all of the women who had the option of choosing CVS outside of the trial opted to make their own choice between CVS and

amniocentesis, rather than being randomised; the authors argued that this emphasised the woman's need to feel in control of her own pregnancy.

Physician preferences

Fahy and Lippman (1988) commented that women who were eligible for the Canadian collaborative trial of CVS versus amniocentesis often cited physician influence as a reason for refusing to participate. The importance of physician influence is commented on by Muggah and colleagues (1987) who undertook a randomised trial in Ottawa of CVS versus amniocentesis. This trial began in April 1985 and continued for 15 months. Difficulty in recruiting patients was reported, one of the major problems being that 307 of the 440 eligible women declined the trial and elected to have amniocentesis. The authors argued that their patients' concern about risk might well vary with the attitude of their physicians towards CVS, concluding that the small number of patients enrolled did not permit any meaningful comparison of amniocentesis with CVS.

In an attempt to measure directly physicians' attitudes to and knowledge of prenatal diagnosis, amniocentesis, CVS, RCTs and, specifically, the Canadian trial, Fahy and Lippman (1988) undertook a questionnaire survey of all registered obstetricians in British Columbia and in Montreal. The response rate was 70%. Most physicians thought prenatal diagnosis was important and that it was their role to discuss and advise their patients on the matter. Physicians were split in their preferences for amniocentesis or CVS (32% versus 34%). According to Fahy and Lippman, physicians who thought CVS was too experimental, who were hesitant about the Canadian trial or who were less likely to discuss the study with patients, were older, less likely to have participated in an RCT previously and less comfortable with randomisation and discussing uncertain risks with patients. Fahy and Lippman (1988) concluded that because physicians acted as gatekeepers, educating them about new technologies and about randomised studies was essential in order to ensure both the patient's access to a new procedure and the success of any planned RCT.

Discussion

In many ways the evaluation of CVS appears to have been exemplary. There was a rapid response to its emergence by national and international bodies. In some countries, randomised comparisons with the standard procedure, second-

trimester amniocentesis, started sufficiently early for the availability of the new procedure to be limited to participants in the RCTs. This was most successful in Canada where it applied to all centres in the country. Although this was attempted in both the MRC European and the Danish trials, the fact that there were other clinical centres nearby offering CVS outside the trial affected recruitment, which eventually led to the Danish trial closing recruitment. It is striking that the Canadian trial started only a year after the first report of CVS performed in a Western country and that the other RCTs started only a year later. In retrospect, the intervening year was one of rapid diffusion of CVS, which is probably why it proved difficult to limit CVS to trial participants in countries other than Canada. The early timing of the trials in relation to publications in Medline are illustrated in *Figure 4.* The peak in publications was seen 5 years later at about the time that the trial results were being reported.

Nevertheless, even as little as 2 years after the first report of CVS and first call for RCTs, it proved impossible to mount such a trial in the USA. The reasons why the attempted RCT in the USA was unsuccessful are uncertain. They are likely to be related to physician and patient attitudes. Certainly, in the MRC European trial, an important reason why women agreed to participate was the support given to the trial by a number of relevant consumer groups, who actively championed it to women. In contrast, in the USA, women were apparently demanding access to CVS and physicians were encouraging them.

Importance of retrospective observational studies

Evaluation of CVS through large-scale trials was not sufficient, however. Because of the very low risk of serious limb-reduction deformities, the trials were not large enough to identify a moderate increase in this respect. By the time that concerns had been raised, CVS had diffused widely and further larger trials would probably not have been possible anyway. Less rigorous, observational studies using the retrospective case–control design were therefore appropriate. (Long-term follow-up of the children from the trials may still be valuable, however, for assessing more common but less serious possible adverse effects, such as more subtle effects on limb development.) The experience of CVS in this respect reinforces the importance of continued, long-term surveillance after the diffusion of a technology.

The finding of a completely unanticipated, serious, long-term hazard also provides empirical support

for the argument of TC Chalmers (1975) that 'randomising the first patient' at least protects half of the participants from what turns out to be the worst treatment. In support of this, I Chalmers has recently challenged the reluctance of some clinicians to randomise early because of the widely held assumption that new treatments are likely to be superior (Chalmers, 1997). In the case of CVS, it can be argued that the women who participated in the RCTs were better off than those who chose CVS outside a trial.

The common argument against starting to randomise very soon after the emergence of a new technology is that the technology may change, making the results of an early assessment irrelevant. The evaluation of CVS illustrates this to some extent. During the recruitment periods of the trials, the transabdominal approach grew in popularity as an alternative to the previous standard, the transcervical method. The Canadian trial, in part because it was the first to get started, only evaluated transcervical CVS and so its generalisability is questioned. Because of the arrival of transabdominal CVS, the protocol of the MRC European trial was modified to allow operators to choose whichever approach they preferred, which made possible some, albeit indirect, comparison between the methods. The results for centres using the transabdominal approach were broadly similar to those of the centres using the transcutaneous method and this increased the study's generalisability. The Danish trial was the most reliable in this respect because it incorporated a three-way randomisation, allowing a direct comparison between the two CVS approaches, as well as with second trimester amniocentesis.

Nevertheless, the finding in the Danish trial that the transabdominal method performed better may be an illustration of another issue in the evaluation of a technology that requires dexterity, that of skill/learning. The principal investigator was one of the first to describe the transabdominal approach and was one of its main proponents; he also performed most of the CVS procedures in the Danish trial. This raises questions about the generalisability of the results of this study to other operators. Attempts to address the issue of skill/learning in the other trials were fairly crude. There was, however, no evidence in the MRC European trial that outcome after CVS improved over the course of the trial. Centres did appear to perform differently (and the trial has been criticised for this) but the numbers in individual centres were too small to assess this reliably and these differences may simply have reflected chance differences between centres.

Although the evaluation of CVS appears exemplary in many respects, were the lessons learnt? In the face of growing concerns about CVS, a new technique, early amniocentesis emerged as a possible replacement for early diagnosis. Despite concerns about possible adverse effects on foetal lung development and foetal viability, early amniocentesis was widely adopted, especially in the USA. Calls for rigorous evaluation through randomised trials largely went unheeded. So far, only two trials have been reported, one only as a letter. The evidence that is available again indicates unexpected risks, again suggesting that the enthusiasm of health technology innovators may be seriously misplaced; HTA may therefore protect both present and future patients.

Chapter 6

Systematic review of diffusion and evaluation of teleradiology and teledermatology, focusing on factors influencing timing

Summary

The purpose of this systematic review is to investigate, within the general field of telemedicine, when and how evaluations of the applications of teleradiology and teledermatology have been undertaken and to clarify the factors that have influenced the timing of such evaluations.

Telemedicine can be defined as the use of transmitted images, voice and other data to permit consultation, education and integration in medicine over a distance. Thus such systems can be used to deliver, for example, radiology and dermatology services to patients who are in a different place from the consultant radiologist or dermatologist. Most telemedicine activity to date has taken place in the USA.

Although telemedicine has been around since the 1950s, the early programmes failed to achieve physician and patient acceptance and were not found to be cost-effective. As a result, when external funding was withdrawn the projects ended and popular interest declined. A cycle of technological development leading to renewed activity, followed by a waning of interest when expectations were not realised, continued approximately every decade. A resurgence of interest has occurred from around 1990 onwards, due to factors such as further technological advances combined with reduced costs, programmes of healthcare reform emphasising the need for improved efficiency, and a demand by rural patients and physicians for equal access to high-quality healthcare irrespective of location.

The development of telemedicine has essentially been technology-driven. Technology providers have been keen to generate new markets for their products by funding telemedicine research and attempting to stimulate both medical and popular interest in such applications.

The potential benefits of telemedicine included immediate access to medical expertise no matter where the patient was, more timely diagnosis and treatment, and the elimination of the need for patients and clinicians to travel long distances between rural areas and urban medical centres. Areas of concern included the threat of malpractice due to misdiagnosis resulting from the use of telemedicine, questions over the acceptability of image quality, and the reluctance of physicians to become involved in telemedicine.

Reports of evaluations of telemedicine systems in the literature have been primarily anecdotal and descriptive. Teledermatology (first referenced in Medline in 1992) seems to be a much more recent application of telemedicine than teleradiology (first reported use in 1950 in the USA). Only one RCT of teleradiology (as indexed by Medline) was identified. Reports in the literature of teleradiology and teledermatology projects generally gave no indication as to the reasons for the timing of the trials, or the details of why they were undertaken at a particular time.

Factors that appear to have influenced the timing of trials of teleradiology and teledermatology include the technology-driven, as opposed to needs-based, nature of telemedicine development. Thus projects have taken place, mainly focusing on technical feasibility, during periods when commercial providers have injected funding. The short time-scale and limited funding of many projects made it difficult to develop and implement RCTs. In addition, telemedicine projects generally served sparsely populated areas, resulting in insufficient patient contacts to provide statistically valid data.

Other factors which continue to affect the spread of telemedicine, and therefore impact on the timing of evaluations, include the threat of litigation and, in the USA, issues of licensing and reimbursement. Damaging malpractice litigation because of misdiagnosis through telemedicine might result in reduced funding and restricted future development. In the USA, physicians are licensed by individual states and doubts over their legal status in other states might act as a barrier to telemedicine expansion. A further potential stumbling block

to widespread implementation of telemedicine in the USA is the lack of reimbursement by the Health Care Financing Administration.

Background

The concept of telemedicine

Telemedicine uses technology comparable to MAS, in that clinical use is made of visual data delivered using remote visualisation technology, that is, video technology. This poses some similar problems in terms of data interpretation and use. The generic technology is also primarily under development in other industries and consequently inaccessible for assessment purposes during periods of its evolution. Telemedicine is also still a fast-changing technology. For these reasons it was selected as one of the six applications examined in this study.

Description of the technology

Merrell (1995) defined telemedicine as the use of transmitted images, voice and other data to permit consultation, education and integration in medicine over a distance, while Perednia and Brown (1995) argued that the defining aspect of telemedicine was the use of electronic signals to move medical information from point A to point B. According to Coles (1995), telemedicine applications could be divided into the three categories of remote diagnosis and consultation, continuing medical education, and medical informatics.

The technology used to undertake telemedicine ranges from telephones to satellites, to state-of-the-art videoconferencing equipment. The fundamental components of a typical telemedicine network, described by Bashshur and colleagues (1975) are:

- the geographic separation of the provider and the client during the clinical encounter or of two or more providers during a consultation
- the use of telecommunication and computer technology to facilitate the interaction between provider and client (or provider and provider) as well as the transfer of information
- appropriate staffing to perform all the necessary functions within such systems
- the development of an organisational structure suitable for implementing telemedicine systems.

Applications of the technology

Coles (1995) argued that the various applications of telemedicine could be categorised into three user groups: store-and-forward, interactive systems, and medical informatics networks. Store-and-forward systems were used to transmit static

images and audiovisual clips to a remote data-storage device, from where they could be accessed by the medical practitioner for review and consultation. Store-and-forward systems had the economic advantage over interactive systems of lower transmission costs, and could be used in applications such as teleradiology, where contact with the patient was not generally required (Coles, 1995).

In a typical interactive system, a coder–decoder (CODEC) transforms analogue images to digital information and compresses the data. At the remote site, another CODEC decompresses the signal and changes it back to the analogue form for viewing on a monitor. A camera and microphone would be located at each of the remote sites and at the hub site. One camera is able to view the patient and radiographer while the other transmits the real-time image (Wright & Loughrey, 1995).

Development and diffusion

The development of telemedicine can be usefully categorised into two time phases: the period from the 1950s until about 1990, and the period from 1990 onwards. According to McLaren and Ball (1995), telemedicine in its most basic form has been around for over 30 years. Coles (1995) stated that the early telemedicine programmes, however, failed to achieve acceptance by patients and physicians and were not found to be cost-effective. When external funding was withdrawn, therefore, continuation of programmes was not viable and popular interest in telemedicine declined.

The American College of Radiology define teleradiology as the electronic transmission of radiological images from one location to another for the purposes of interpretation and/or consultation. Teleradiology is by no means a recent development; the first recorded instance was in 1950 (Gershon-Cohen & Cooley, 1950).

The impact of hardware advances on adoption

McLaren and Ball (1995) pointed out that since the introduction of the early telemedicine programmes, each decade had witnessed a resurgence of activity as new ways of generating images or transmitting data were developed. They noted, however, that in most cases once the technical feasibility was demonstrated, descriptive reports were written and recommendations made for further research, which in fact rarely materialised. This cycle would then be repeated as new technology was developed, which offered in turn even faster transmission of yet higher quality images (McLaren & Ball, 1995). Expending time

and resources on systematic evaluation was less attractive than trying out the next generation of equipment. Both the limitations of available technology and suboptimal use of existing technology contributed to unsatisfactory results.

An editorial in *The Lancet* (14 January 1995) commented that the recent resurgence of interest had yet to have a major impact on mainstream medical services, while Perednia and Brown (1995) argued that the recent growth of clinical telemedicine was now creating an increasing demand for information about its safety, effectiveness and clinical utility. Bashshur (1995) argued that knowledge of telemedicine's true effects lagged far behind the current rush to establish telemedicine systems. He claimed that serious research and evaluation were still in a phase of gestation, and that it would be some time before hard data were available to assist policy making in this area (Bashshur, 1995).

The early phase

Coles (1995) commented that the technology used in the implementation of the original telemedicine programmes was often inappropriate, in ways varying from the under-utilisation of available resources to the adoption of expensive interactive, broadband networks which exceeded operational needs. In both cases, inappropriate usage rendered the pilot projects economically impracticable. In addition, the technology was constrained in terms of transmission-relay, audio quality, and equipment reliability (Bashshur, 1995).

In the USA, the National Aeronautics and Space Administration played an important part in the early development of telemedicine, providing much of the technology and funding for early demonstration projects (Bashshur & Lovett, 1977). According to Grigsby (1995), the long-distance telecommunications networks then available could not support the wide bandwidth transmission necessary for real-time videoconferencing. Computing resources were costly and limited in speed, power and flexibility so that, as with earlier programmes, when funding ran out the projects could not be sustained.

Maclean (1996) argued that throughout the 1970s adventurous technology-driven telemedicine projects resulted in the creation of more problems than solutions. Difficulties included unexpectedly high costs, inadequate training for remote practitioners and unfulfilled expectations of the health professionals involved. As a result of these difficulties, from about 1978 until the

mid-1980s, few studies were undertaken on telemedicine. Coles (1995) pointed out that, with the exception of a 20-year-old telemedicine programme in Newfoundland, none of the projects implemented before 1986 had survived beyond their original grant-funding cycle.

The second phase

Technological advances along with reduced set-up and transmission costs saw a renewal of interest in telemedicine in the late 1980s (Maclean, 1996), with programmes implemented in Canada, Scandinavia and Europe (Coles, 1995). Nevertheless, by 1990 in the USA there were still fewer than ten telemedicine programmes in existence (Perednia & Brown, 1995).

From about 1990, a resurgence of interest in telemedicine took place (Maclean, 1996). According to Grigsby (1995), this renewed interest was caused by the convergence of a number of economic and technological factors, including the development of high-resolution video cameras and monitors, digitisation and compression of data, the development of powerful, inexpensive and easy-to-use computing resources, improvements in the quality and distribution of telecommunications networks and applications, widespread interest in controlling healthcare expenditures and in reorganising the ways in which most people received medical care, and economic and political conditions which encouraged capital investment in this sector. Coles (1995) also pointed to an increasing demand for universal access to high-quality medical care, irrespective of location. She claimed that, as a result of such developments, telemedicine had now become a workable reality, noting that in 1995 an estimated 110 active telemedicine programmes were recorded in the USA, together with significant activity in Europe, the Middle East, and Japan. According to Smits and Baum (1995), however, telemedicine is still in many ways in its infancy. A search of the Telemedicine Information Exchange on the World Wide Web undertaken in February 1996 identified 45 active programmes concerned with teleradiology, of which 42 were based in the USA. A similar search on teledermatology identified six programmes, five of which were based in the USA.

Objectives

The objectives of the review were to identify the potential benefits and disadvantages of telemedicine, to discover when and how evaluations of teleradiology and teledermatology had been

undertaken and to clarify the factors that influenced the timing of those evaluations including commercial pressures.

Search strategy

Literature searches were carried out primarily on the databases Medline, Embase, and the Cochrane Database of Systematic Reviews. The Telemedicine Information Exchange on the World Wide Web was also consulted. The database searches were supplemented with articles identified by other means, including contact with experts in the field, hand-searching of reference lists of key articles obtained, and monitoring the contents of a small number of medical journals in 1996 (see Appendix 1 for a more detailed description of the search strategy, criteria for inclusion, and methods).

Studies included

Potential benefits of telemedicine

One of the advantages of telemedicine, according to Coles (1995), was the potential for universal access to high-quality medical care, irrespective of geographical location. There were four key classes of potential beneficiaries: populations with limited access to health-care; residents of rural and remote areas; other medically under-served demographic populations, such as inner city communities; and finally, situations where there were inequalities in the geographic distribution of healthcare providers.

Bergman (1993) argued that, in many cases, telemedicine provided medical care to communities that did not have medical experts and speciality services. The benefits included immediate access to medical expertise, no matter where the patient was, more timely diagnosis and treatment, and the elimination of the need for patients to travel long distances from rural areas to urban medical centres.

Rinde and colleagues (1993) commented that remote hospitals were able to gain access to new services and specialist competence, and could obtain consultations without sending patients to regional hospitals. If difficulties arose, the remote physicians could receive almost immediate assistance from their specialist colleagues. By this means, telemedicine could break through the professional isolation otherwise associated with geographical separation. However, all these potential benefits depend on an evaluation of the technology

confirming that it does, indeed, confer net benefits on these groups.

Potential disadvantages of telemedicine

By its very nature, widespread adoption of telemedicine could impact on many aspects of the healthcare delivery system. A number of areas of concern and potential drawbacks to the wider implementation of telemedicine have been raised in the literature, including legal and funding issues, image and, hence, diagnostic quality, and user acceptance in general. An editorial in *The Lancet* (14 January 1995) drew attention to a few of these possible drawbacks. There was some suggestion that diagnoses based on transmitted images might be less accurate than those based on the originals. Also, patients might increasingly demand specialist consultation, such that the telemedicine programme might not be able to satisfy the demand. In addition, specialist telemedicine centres would probably be expected to provide a fast service, as undue delays would question the viability of the programme compared with alternative options.

McLaren and Ball (1995) argued that although telemedicine had been presented as offering both education and service delivery to isolated practitioners, these aims might conflict where a telemedicine provider sought to increase demand for remote diagnostic services. In such a scenario, remote practitioners might be less inclined to develop their own skills, resulting in a shift towards high-cost diagnostic practice. Coles (1995) argued that the technology-driven nature of telemedicine development had created problems of user-functionality, and that changes in approach were needed to ensure that the system was designed for the needs of the patient–physician interface. Another drawback, according to Crump and Pfeil (1995), was that the normal telemedicine consultation was more time-consuming than the equivalent on-site encounter.

To allow fast transmission of images, Wright and Loughrey (1995) pointed out, the data for digital images had first to be compressed. In general terms, greater compression resulted in faster transmission but also led to a higher risk of data loss. Computed tomograms and magnetic resonance images were readily compressible because they remained of diagnostic quality even when reconstructed with relatively low resolution. However, standard plain film radiographs could tolerate much less compression if they were to remain of diagnostic quality, and Wright and Loughrey commented that radiologists remained divided on whether the detail of the received image was

of an acceptable standard. An article in the *Electronic Engineering Times* (29 April 1996) claimed that the biggest technical obstacle to widespread expansion was the lack of a fully functional broadband network which could transmit voice, video and data at an affordable cost. Thus, there are still questions as to whether the technology and the infrastructure are sufficiently well-developed to support widespread adoption.

There is also the question of 'who will pay?'. With specific reference to the USA, Bergman (1993) argued that one of the current obstacles to widespread implementation of telemedicine was the lack of reimbursement by the Health Care Financing Administration on a national level. This, along with physician reluctance to use telemedicine systems, was seen by Grigsby (1995) as being responsible for the low volume of telemedicine consultations associated with most of the existing programmes.

Another significant potential issue is the threat of malpractice litigation resulting from misdiagnosis arising from the use of the telemedicine system. A related issue was the legal status of the specialist consultant in a remote state (in the USA) or remote country, in the case of international programmes. There may also be differences in the way in which the law might apply to material in paper and computerised formats. In addition, as more specialists, physicians and other health professionals collaborate in telemedicine programmes involving computerised systems, so the potential for misuse of data and breaches of data security increases, as does the ever-present risk of system breakdown.

The question of potential costs has also been discussed in terms of net benefits received. Physician reluctance to become involved with telemedicine might be due to one of the principal drawbacks to its wider implementation, namely the lack of formal studies of cost-effectiveness (Grigsby, 1995).

Media reports

Papers retrieved by the Promt search tended to concentrate on the telemedicine market and the opportunities perceived by the commercial technology providers. For example, a report in the *Electronic Engineering Times* (29 April 1996) forecast that telecommunications carrier revenue from telemedicine would surge between 1998 and 2001, and that the biggest technical roadblock was the lack of a fully functional broadband network that could transmit voice, video and

data at an affordable cost. Other reports described a proposed telemedicine system in Singapore, discussed the potential of telemedicine to help solve healthcare problems in the USA, considered the potential of telemedicine to generate profits for technology providers, reported a telemedicine trial in Hawaii, and commented on expansion into teleradiology services by telecommunications companies.

The TeleMed 96 international telemedicine conference took place in London in November 1996. Following on from the conference, Illman reported in *The Guardian*, 26 November 1996, that long-distance medical treatment using video technology and digital communications was becoming a reality for people who could not reach a doctor. He noted that, although the ultimate 'remote medicine' occurred in outer space, telemedicine would also transform everyday medical practice. A number of test applications of telemedicine were described, such as on aircraft, ships and off-shore installations, for use in medical training, to aid the elderly at home, and for remote diagnosis of heart scans. The article also referred to the savings made by a nurse-run minor injuries clinic in London which had a video link to a Belfast hospital with access to specialist doctors, thus removing the need to employ doctors directly. The overall tone of the article regarding the potential of telemedicine was very positive, which would have been conveyed to the general reader.

An article in *The Economist* (11–17 January 1997) described a telemedicine service based at Hays Medical Centre, Kansas, USA, that provided a consultation service to a vast area of the western Kansas plains. The telenursing part of the service allowed more patients to be seen in the time available, with substantial cost savings for televisits compared with home visits. The article noted that telemedicine offered huge potential savings, many people in rural America looking upon it as a miracle cure.

Attention was drawn, however, to the problem of reimbursement. The rules governing payment for services performed under Medicare, the US healthcare system, meant that the only applications of telemedicine that currently qualified for reimbursement were teleradiology and telepathology. The article noted that the US Health Care Financing Administration had only recently begun a 3-year trial of Medicare reimbursement for teleconsultations in the states of Georgia, Iowa, North Carolina and West Virginia. It is expected to be about 5 years before any recommendations

arising from the trial are assessed and turned into law. This cautious approach is being adopted because the federal government wants to know in advance how much telemedicine is likely to cost, once it is sanctioned under Medicare; if access to expert consultation became much easier through telemedicine, then the overall volume of service might rise. This situation demonstrates how central government, through withholding reimbursement, can delay the widespread diffusion of a health technology into routine practice, and can ensure that adequate evaluation is undertaken before such diffusion is allowed to proceed.

Impact of on-going and new technological developments

The driving force behind telemedicine developments has not been limited solely to advances in communications technologies. *Table 9* demonstrates the way in which major technological advances and inventions which have taken place in the socio-political field, and in the telecommunications,

TABLE 9 *Telemedicine – chronology of relevant technological advances made outside medicine*

Decade	Socio-political	Telecommuni-cations	Micro-electronics	Video	Other
1940s		• Co-axial cable developed (transmission distance)		• Black and white television	
1950s	• USA Navy first bounced radio signals off the moon • Start of Cold War	• Telex arrived 1956 Canada 1957 USA (66 words/minute = 50 bps)	• Signal boosters invented (signal strength)		
1960s	• Space Race • First satellites 1960 Echo 1962 Telstar	• Digital technology beginning to replace analogue technology		• Colour television	
1970s		• First double polarisation (multiple channel capacity) satellite 1976 Satcom	• Large Scale Integration (LSI) technology invented (echo suppression)	• First video-telephone 1970 Bell System (maximum distance 6 miles)	• Advances in battery technology (no longer do satellites have two 44-day down times/year)
1980s	• Global satellite networks 1980 Ariane 1982 Delta 1984 Titan	• First ISDN public service 1989 Singapore	• Personal computer 'arrived' 1981 (ZX81)	• Pulse code modulation (PCM) developed (detects only image movement)	
1990s	• De-regulation of much of world communications networks • Spiralling demand for telecommuni-cations capabilities • The 'global marketplace'	• Widespread ISDN • 'Functioning' multi-channel ISDN • Arrival of WAN, MAN, LAN, with bandwidths up to 50 Mbps	• PC developments (faster, smaller and cheaper)	• PC processing technology (video-conferencing increasingly available, smaller units, and cost down from £35,000/unit to £2500/unit in 5 years)	
Future	?? Further develop-ments of NHS purchaser-provider dynamic	?? Further world standards agreed for telecommunication technologies (H320)	?? 'Feely' glove ?? Wrist watch videoconferencing	?? HDTV quality videoconferencing	?? New video compression algorithms

micro-electronics and video industries have allowed physicians to follow in their wake. To this extent, it has not been the healthcare professionals who have led the way in deciding when to evaluate telemedicine but rather their enthusiasm has been stimulated by external factors. For example, in the socio-political field, the 1960s 'Space Race' between the USA and the USSR, when both superpowers were vying to be the first to land an astronaut on the moon, proved to be a major source of funding and support for many early telemedicine demonstration projects, such as STARPAHC (Space Technology Applied to Rural Papago Advanced Health Care) (Fuchs, 1979).

In the field of telecommunications, a significant milestone was the arrival in the public domain of the ISDN (integrated services digital network) in Singapore in December 1989. This enabled users for the first time to access on demand, multi-bandwidth, pay-only-as-you-use videoconferencing (Michaelis, 1990) and was, without doubt, the catalyst that prompted the surge in interest in the development, and subsequently evaluation, of telemedicine in the 1990s.

Can information such as that provided in *Table 9* allow for the prediction of future 'leaps forward' in the potential uses of interactive video for health-care? For example, it can safely be speculated that new video compression algorithms will arrive during the next decade. Such a development in the field of mathematics will have a significant impact on future telemedicine developments, by virtue of the fact that improved video compression will result in better picture quality in relation to the available bandwidth, in turn leading to enhanced image quality in videoconferencing. This development, it could be argued, might be the catalyst for many of the healthcare professionals who are currently sceptical of telemedicine to become more open in their judgement.

Impact of commercially-motivated technology providers

The main driving force in telemedicine appears to have been developments in communications technology (McLaren & Ball, 1995; Maclean, 1996). Coles (1995) argued that the impetus for the current resurgence of telemedicine was due, at least in part, to the self-interest of technology providers, who were keen to expand their traditional customer base and open up new markets. By funding telemedicine research in medical organisations and academic institutions, telecommunications companies had succeeded in stimulating interest among the medical community and the general public alike. This, combined with an increasing demand for high-quality health-care created by rural and chronically underserved populations had helped to generate new markets for their products (Coles, 1995).

The source for much of the following information on telemedicine technology providers is the monograph by Coles (1995). Kodak was among the first of the conventional imaging product manufacturers to recognise the opportunities presented by the changing market and the substitution of hard copy media with digital equivalents; the company has been a strong supporter of telemedicine and has funded considerable research into potential markets. This development of general software applications tailored to meet the complete needs of customers operating within particular sectors has made the healthcare industry an attractive market for software providers. The exceptional growth of Microsoft, the world's leading independent software supplier, has ensured it a powerful position in the development of telehealth and on-line healthcare networks (Coles, 1995).

The resurgence of interest in telemedicine has also led to the development of providers who specialise in offering independent solutions to meet their customers' requirements. WellCare is a prime example of these new telehealth solutions providers, formed in 1992 to take advantage of the opportunities created by the commercial demand for telemedicine services. WellCare originally focused on the commercial provision of global teleradiology services. The success of WellCare can be attributed at least in part to the company policy of using existing technology to provide immediate solutions, rather than concentrating on programmes concerned with technological development (Coles, 1995).

AT&T is the largest provider of telecommunication services in the USA. The company has funded and participated in an extensive programme of telemedicine activity aimed at encouraging the widespread introduction of such systems. AT&T began marketing the Picasso® still-image phone in 1995 as a diagnostic tool which would enable physicians to send high-quality still images over standard telephone lines, while simultaneously conducting a real-time teleconsultation with a remote physician. The Picasso phone has been successfully trialled in a variety of medical fields, including dermatology (Coles, 1995). *The Cable Act 1984* resulted in the

division of part of AT&T into local telephone carriers, with the regional franchises allocated to Ameritech, Bell Atlantic, Bell South, Nynex, Pacific Telesis, SBC Communications and US West. In addition to the telemedicine activities undertaken by AT&T, these local operators have also become involved in funding regional telemedicine programmes (Coles, 1995).

Another major USA-based telecommunications organisation, GTE, has placed ISDN technology at the centre of its telemedicine initiative, allowing simultaneous voice and data transmission across the same line. The GTE telemedicine initiative has concentrated on a number of principal applications, including teleradiology and remote diagnosis and consultation. Sprint is predominantly a long-distance telecommunications carrier operating throughout the USA. The organisation has made inroads into a number of telehealth areas. Within Sprint's telemedicine initiative, considerable emphasis has been given to the development and implementation of teleradiology systems (Coles, 1995). Bergman (1993) drew attention to the creation of some businesses specialising in teleradiology services, such as Teleradiology Associates, who interpret X-rays and other images for hospitals nationwide in the USA.

In the UK, BT (British Telecommunications), in response to a growing demand for telemedicine systems, has developed the CARE project, initiating a series of telemedicine trials designed to gain an insight into the potential impact of telehealth services (Coles, 1995).

Coles (1995) argued that telemedicine, with its demand for high bandwidth for the transmission of high-quality medical images, provided the telecommunications industry and technology providers with an expanding base of additional healthcare customers to supplement the existing markets. Further dramatic growth seems probable. An article in *Electronic Engineering Times* (29 April 1996) forecast that revenue from telemedicine was expected to surge between 1998 and 2001. The Insight Research Corporation of Livingston, New Jersey, USA, estimated that health information networks would generate almost $2 billion in telecommunications carrier revenue alone by the year 2001, not including the revenue from equipment and software needed for telemedicine applications. Thus there has been a strong profit-oriented technology push from manufacturers, although clearly there has also been a degree of receptiveness on the part of healthcare providers.

Evaluation of telemedicine

There is growing awareness of the need for proper evaluation of telemedicine applications. McLaren and Ball (1995) reported that, despite rapid technological advances, evaluations of telemedicine systems had nevertheless been largely superficial. Also, as Tangalos (1995) stated, although telemedicine had been in existence for a good number of years, most published accounts of such projects tended to be descriptive; proper clinical trials were necessary to validate the claims for telemedicine systems which were being made by their current proponents. The view that most recent telemedicine reports were simply descriptive, with little precise reporting of the technology used, was echoed by Crump and Pfeil (1995); most recently published studies were small and, typically, included fewer than ten patients. They argued that the issue of the cost-effectiveness of telemedicine had to be addressed, with analyses including not only hardware and operating costs but also the costs to patients and physicians in terms of time and transport. The dearth of systematic empirical research regarding the true effects of telemedicine on the cost, quality and accessibility of care had left the field ripe for speculation and opinion (Bashshur, 1995).

The bulk of studies on telemedicine which had been undertaken were confined to North America, with only a minority taking place in Europe or elsewhere (Maclean, 1996). He also noted that almost without exception these studies were pilot projects, written up in the clinical literature as anecdotal reports.

A recent attempt to review the technology systematically found the body of data available to be unsatisfactory. Coles (1995) reported that, in 1993, the American Agency for Healthcare Policy and Research commissioned a study investigating the potential benefits of telemedicine in the USA. The primary aim was to establish whether the provision of health services using telemedicine was medically safe and effective. The first phase of the project was a comprehensive review of the literature on telemedicine, which found little material on the effectiveness of telemedicine. The second phase included a series of case reviews of established telemedicine programmes throughout the USA to evaluate the current status of telemedicine. The third phase was a study of quality assurance, utilisation review and payment policy. The extension of Medicare cover to include telemedicine services was rejected by the authors of the study on the grounds of the experimental nature of the applications and the lack of methodologically sound research (Coles, 1995).

Evaluation of teleradiology
Hardware aspects

Initial research, using satellite as the transfer medium, was found to be both costly and technically demanding. Technological developments have since made possible the transfer of images along telephone lines, and hardware and software costs have also fallen (Wright & Loughrey, 1995). Although interactive videoconferencing systems have received the most media attention, Wright and Loughrey (1995) argued that many supporters of teleradiology believed that the simpler, less expensive store-and-forward systems would become the most popular choice.

Coles (1995) noted that with the recent proliferation of teleradiology programmes, concern had been expressed about the integrity of digitised images and their suitability for diagnostic evaluation. Compressing the image is essential for the rapid exchange of data across a network and the greater the degree of compression, the faster the transmission of the signal. The drawback is that as the rate of compression increased, so too does the risk of data loss, possibly compromising the diagnostic accuracy of the image received by the remote radiologist. Although computed tomograms and magnetic resonance images have been shown to be suitable for compression, Coles (1995) argues that standard plain film radiographs could tolerate a far lower ratio of compression before the reconstructed image loses the required quality of resolution.

The American College of Radiology is the principal professional body of radiologists, radiation oncologists and clinical medical physicists in the USA. In 1994, the College introduced a standard for teleradiology, for use by practitioners undertaking the remote diagnosis of digitised medical images, which defined goals, qualifications of personnel, equipment guidelines, licensing, credentialling, liability, communication, quality control, and quality improvement for teleradiology. The introduction of these guidelines had made a positive contribution to the success of teleradiology and its achievement of reimbursable status in the eyes of the Health Care Financing Administration (Coles, 1995). However, Grigsby (1995) maintained that professional opinion was still somewhat divided over the effectiveness of teleradiology, although acknowledging that its relatively widespread use might be interpreted as a positive sign.

Descriptive studies

Information on a few of the many descriptive studies of teleradiology to be found in the clinical literature is presented here.

Gayler and colleagues (1981) described a laboratory evaluation of a microcomputer-based teleradiology system, where 100 radiographic examinations in total were interpreted by 12 radiologists. Each radiologist interpreted 50 cases, 25 in film mode and 25 in radiology mode. The trial took place during July and August, 1980. The radiologists' scores for findings, impressions and confidence levels were significantly lower for radiological images viewed on the teleradiology system; however, Gayler and colleagues argued that the quality of the images provided by the teleradiology system was high enough to warrant further study. No information was given as to why the trial was undertaken at that particular time, other than a general statement that the lack of timely access to radiology services was a major problem for millions of Americans in rural areas and that increasing attention was being focused on the feasibility of using teleradiology to transmit radiographic images over telephone lines from underserved areas to consulting radiologists.

Gitlin (1986) described two American field trials undertaken in 1982 and 1984 involving the armed forces, with five medical care facilities participating. The trials were undertaken in an effort to improve radiological services to personnel on small bases and ships at sea. It was also envisaged that the successful development of a teleradiology system not only would serve military personnel during peacetime but would also improve medical care under combat conditions. The 1982 trial results indicated that the interpretation of the radiographic films was more accurate than the teleradiology system but the results were sufficiently encouraging to warrant a similar evaluation of a system that would provide higher resolution. The 1984 field trial again reported superior accuracy of radiographic film interpretations but it was argued that the teleradiology concept continued to show promise as a method of providing radiographic services to certain populations.

Rinde and colleagues (1993) gave details of an on-going telemedicine project in rural North Norway. Since 1988, Norwegian Telecom Research, in collaboration with the University Hospital of Tromsø, has been conducting a project in which remote diagnoses have been made. The main challenge was seen as giving patients access to special expertise and medical services in a way which was practical, economical and reduced the need to move patients away from their home districts. An additional aim, by providing access to a support network, was to make the health professions in rural districts more attractive.

According to Rinde and colleagues (1993), the University Hospital of Tromsø has, during the last few years, gained experience with remote diagnosis in a number of medical disciplines, including radiology and dermatology. In radiology, trials were conducted on expert consultation involving the use of videoconferencing facilities and high-quality image transfer between workstations. The image quality in videoconferencing was not good enough for displaying all details and grey levels, and in order to get diagnostic quality for a routine teleradiological service it was necessary to use high-quality digitisers and monitors. The system involved daily scanning of analogue films, transmission of the digital images to the hospital, diagnostic examination on a multiscreen workstation, and digital transmission of dictated reports back to the local clinic.

Rinde and colleagues (1993) concluded that the results so far had shown that teleradiology might represent a viable solution for small clinics lacking qualified radiologists, as well as for clinics with a single radiologist who needed ready access to colleagues. They claimed that the project had also demonstrated that teleradiology could now be used routinely to serve local clinics, with little loss of picture quality and substantial gains in the quality of care. However, McLaren and Ball (1995) argued that although the North Norway project had run successful pilot studies of remote diagnosis in dermatology, radiology and other disciplines, the impact of telemedicine on the health of the population was nevertheless still uncertain.

RCT of teleradiology

The literature searches undertaken identified only one trial of teleradiology which was classified as an RCT (in Medline). The purpose of this study by Scott and colleagues (1993) was to evaluate whether radiologists performed equally well with plain radiographs or digitised images displayed on a video monitor in the interpretation of difficult orthopaedic trauma cases. Interpretations from film and those made from a teleradiology system with a spatial resolution of 2.35 line pairs per mm were compared in 120 difficult cases (60 with the selected abnormality and 60 control cases). Seven senior radiology residents and one radiology fellow each interpreted 60 randomly ordered cases with the teleradiology system (1280 x 1024 pixel monitors) and 60 cases with the original radiographs. The hypothesis to be tested was that radiologists performed equally well with film images or images on the teleradiology screen.

Full control of trials was not achieved because of the rapid change in available technology over the period of the trials. In 1989, when the teleradiology trial was planned, the system to be evaluated was considered 'state of the art'. It comprised a 0.210 mm spot-size laser digitiser and a 1024 x 1280 pixel monitor image display workstation. The 1989 capital equipment and annual operations and maintenance costs of the teleradiology system used were approximately $352,000 and $22,000, respectively (including transmission and receiving sites). Scott and colleagues (1993) reported that by July 1992, a state of the art teleradiology system for the interpretation of bone fractures could include a laser digitiser with 0.105 mm spot size and 2048 x 2560 line display monitors. The cost of such a system in 1992 was almost equivalent to the cost of the original 1989 system, so that over a period of 3 years the available spatial resolution of the digitiser and display monitors had doubled while the cost had remained fairly constant.

Although evaluation of the time required for radiographic interpretation was not a primary objective of the study, Scott and colleagues (1993) pointed out that the cost-effective use of a radiologist's time was, nevertheless, an important practical consideration. They found that, on average, interpretations with on-screen images required 1.54 times longer than interpretations with film. Because of the cumbersome nature of the image manipulation features, participants often failed to take full advantage of the contrast, brightness, and magnification functions in all cases, which thus placed interpretations with on-screen images at a disadvantage.

Scott and colleagues explained that the case-mix in the study was deliberately skewed towards more difficult than average fracture cases, which would dramatically accentuate any potential deficiencies of the teleradiology system. They claimed that their results indicated a need for special care in excluding the possibility of fracture, when using a teleradiology system with similar specifications to the one used in the study.

The overall conclusion reached was that the teleradiology system tested was unsatisfactory for primary diagnosis of subtle orthopaedic images. They acknowledged that lack of training and experience of the system, which contributed to the poor results, might be related to a learning curve associated with the use of digital display workstations (Scott *et al*, 1993). The published report of the trial gave no specific details about when it took place, how long it lasted, or the why it was undertaken at that time. It does, however,

illustrate the difficulties of achieving satisfactory RCTs with fast-changing technologies.

Evaluation of teledermatology

The literature searches identified few references on teledermatology but none that contained reports of RCTs.

The rationale for adopting teledermatology

According to Coles (1995), skin disease accounts for more than 2% of total global healthcare expenditure. In the USA, approximately 80 million people (30% of the population) have at least one clinically significant skin condition. The incidence of skin cancer, the most common dermatological malignancy, is also continuing to rise around the world.

Coles (1995) noted that specialised dermatological care was traditionally based in major urban centres. The fact that many remote communities lacked access to such facilities rendered them more susceptible to skin-related diseases. Local practitioners did not possess the same level of expertise as their specialist colleagues to provide accurate diagnoses and the alternative was to send rural patients on long journeys to specialist urban centres for examination.

Coles (1995) stated that teledermatology represented a practical solution to the problems associated with the unequal geographical distribution of dermatologists. She argued that teledermatology could be used to bridge the gap between urban specialists and their rural counterparts by providing diagnostic expertise both to meet consultation requirements and to increase rural physicians' knowledge through expanding their practical experience of dermatological care.

Perednia and Brown (1995) described the Oregon High Performance Computing and Communications Initiative. They explained that, as there were only two dermatologists in the entire eastern two-thirds of Oregon, for many rural residents this meant that the nearest dermatologist was hundreds of miles away. This teledermatology project was designed to bridge this healthcare gap through the innovative use of a store-and-forward teledermatology system, using high-speed computers, moderately high-resolution colour cameras, wide-area networks and full-colour digital image storage. The system was to be installed in three primary care clinics serving rural, dermatologically underserved areas of Oregon. The images were to be transmitted from the rural centre to a consulting dermatologist at Oregon Health Sciences University for review and discussion with the primary care physician. The system was store-and-forward rather than interactive, so there was minimal disruption of the normal clinical routine at either end. The project was expected to demonstrate how teledermatology could enhance the provision of dermatological care to rural areas of the USA (Perednia & Brown, 1995).

Plans for evaluation

Perednia and Brown (1995) commented that very little was known about dermatologists' ability to diagnose skin conditions accurately using electronic images. The basic research component of the project was designed to verify the ability to make accurate diagnoses based upon digital images, and define the minimum technical specifications needed to ensure that diagnostically important information was captured. The final result was to be a low-cost, store-and-forward teledermatology system that would be subjected to field testing in the clinical phase of the project. This would be undertaken in a series of steps. From the data collected, a prototype teledermatology imaging and transmission system would be assembled that had been optimised to be maximally informative, easy to use and relatively inexpensive to assemble.

Perednia and Brown (1995) explained that the utility of the teledermatology system developed in the basic research phase of the project would then be evaluated in a 2-year clinical trial. The purpose would be to establish whether this technology would improve the process of healthcare delivery by increasing information flow and reducing isolation; improve the provision of dermatological care; and increase the primary care provider's knowledge of dermatology. As in the basic research phase, the clinical evaluation would proceed along several fronts for the duration of the project. One important goal would be to increase the overall ability of primary-care physicians to recognise skin cancers and other important conditions.

The report of the Oregon project gave no information about when the project was to be initiated or the expected numbers of patients. Other than a general statement that the growth of clinical telemedicine was creating a demand for information about the safety, effectiveness, and clinical utility of such technologies, no reasons were given for the timing of the trial.

Rinde and colleagues (1993) gave details of the teledermatology component of the Telemedicine in North Norway project. Before the project began,

dermatologists in northern Norway provided an outpatient service to small hospitals and health centres. In the trial starting in 1989, this service was replaced by a videoconferencing link with the University Hospital of Tromsø. The general practitioner's patients were brought to a video-conferencing studio at a local centre twice a month. A dermatologist in the videoconferencing studio at the University Hospital then received an account of the patient's condition from both patient and doctor, and the camera was focused on the area concerned. The dermatologist was able to view either a live image or a high-quality still image. A diagnosis would then be arrived at collaboratively by the general practitioner (GP) and the specialist dermatologist. One benefit of the system, in addition to the dermatologist's diagnosis, was an expected improvement in the GP's knowledge of skin diseases, which contrasts with the prediction by McLaren and Ball (1995) that a GP would make less effort to acquire expertise if telemedicine gave easy access to a specialist. The system is now in routine use and is to be extended to other remote locations (Perednia & Brown, 1995).

As in the report of the Oregon project, no specific information was given as to why the teledermatology element of the telemedicine in the North Norway project was initiated in 1989, other than a general statement that new ways were being sought of using telecommunications to provide health services in rural areas which were equivalent to those available elsewhere.

Economic aspects

The costs of introducing and using telemedicine have been met from a number of sources. Coles (1995) pointed out that, owing to the increasing demand for high-quality health-care from rural and under-served populations, commercial providers had sought to stimulate interest in telemedicine at both medical and popular levels, through funding programmes directed at hospitals, physicians and patient populations. Technology companies saw an attractive potential market in the provision of two-way interactive video systems and the transmission of high-quality medical images using expensive broadband digital bandwidth. They had already made major investments in the generic technology. To use it in medical applications was a relatively small additional cost. For this reason the medical products did not have to carry the full cost of developing the technologies.

Coles (1995) argued that as the average telemedicine project required only intermittent use of high-capacity data systems, the expense of digital transmission was generally unjustifiable. She pointed out that the developed world was already connected by conventional telephone lines, and data compression and decompression technology had enabled the utilisation of this existing network. The on-going expenditure on infra-structure for other applications was creating the possibility of high-quality images in telemedicine without additional investment.

Crump and Pfeil (1995) pointed out that the cost of telemedicine technology was rapidly decreasing. CODEC units that cost more than $100,000 just a few years ago, now cost about $20,000. Video-conferencing equipment (television monitor, remote-controlled pan-and-tilt video camera, CODEC, microphone, video recorder, and cabinet) was available for $50,000–60,000, while store-and-forward systems using slow-scan processors could be obtained for about $20,000.

Although continuing technological developments were seen as being partly responsible for the increase in overall healthcare costs, telemedicine was nevertheless widely expected to lead to reduced costs (Coles, 1995). Also, advocates of telemedicine (Bergman, 1993) argued that costs should decrease because fewer specialists would be needed and there would be more reliance on nurse practitioners and physician assistants. The critics' response, however, was that there was no proof that this would occur.

In the USA, some telemedicine demonstration projects had received state funding, while others had received assistance from the Rural Health Policy Office and the Federal Rural Electrification Administration (Bergman, 1993). In the case of the Western New York Project, in which the Erie County Medical Centre in Buffalo, New York, was connected to four rural sites, equipment and telecommunications vendors had provided initial funding for the project; however, the project was now seeking further assistance to help pay for videoconferencing and teleradiology equipment. Thus equipment and project costs were being paid for by the taxpayer but not from medical budgets, while development and infrastructure costs were met by other budgets (Bergman, 1993).

Publication trends

The number of references on teleradiology and teledermatology identified through a search of the Medline database are shown by year in *Figures 13* and *14,* respectively. In the case of teleradiology, the annual number of references remained variable and relatively low until
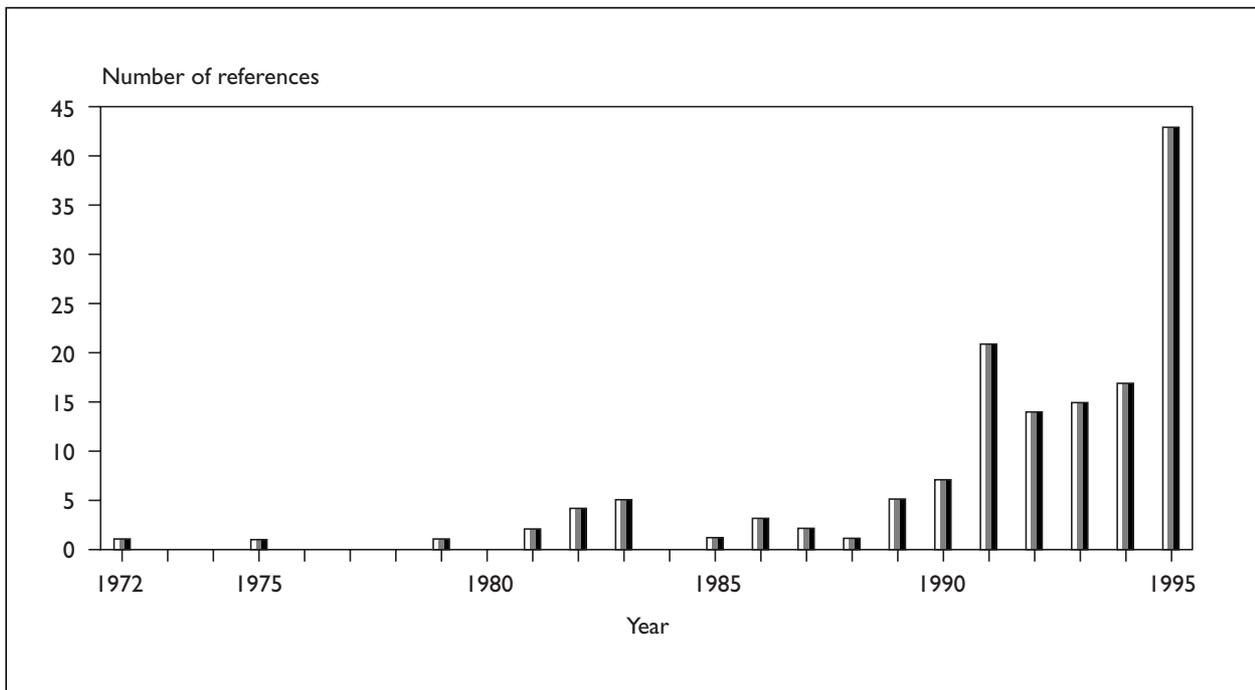
FIGURE 13 *Medline references – teleradiology*
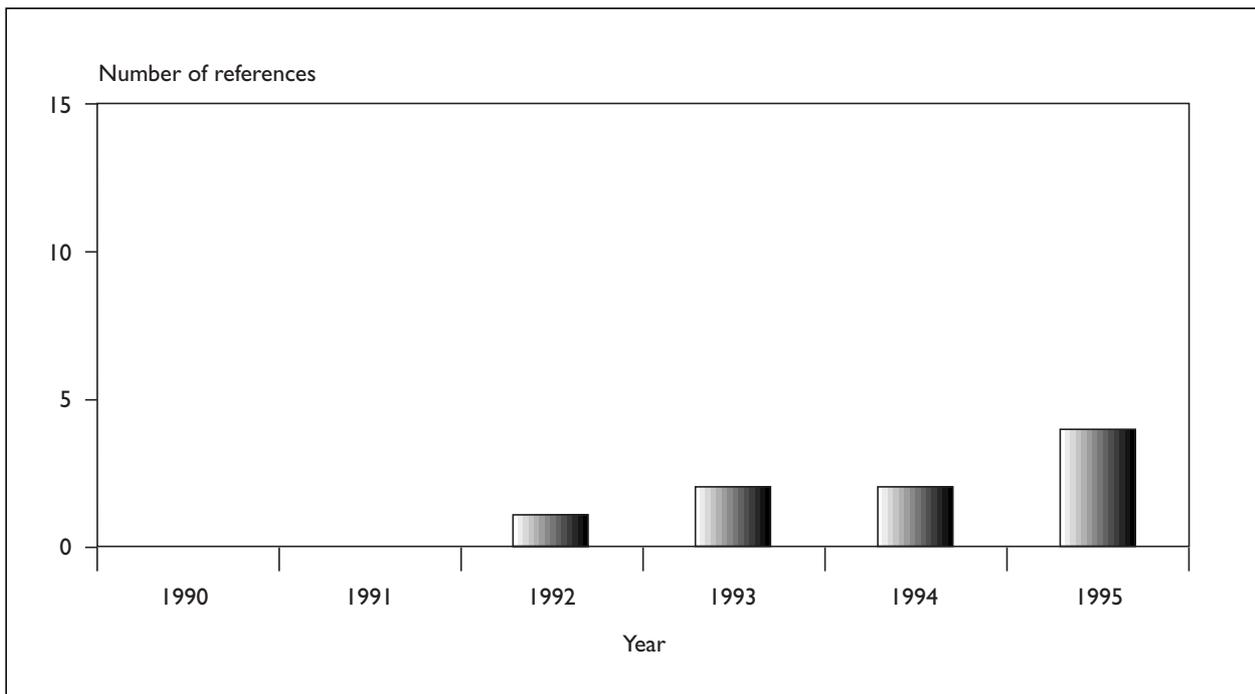
FIGURE 14 *Medline references – teledermatology*

about 1990, with a generally increasing trend discernible from then onwards, culminating in a major increase to 43 references in 1995. This reflects the cyclical pattern of telemedicine development referred to earlier, whereby approximately every decade technological developments generate new activity which subsides when funding runs out and expected benefits do not materialise.

The upward trend from 1990 reflects the current resurgence of interest in telemedicine. The total number of references is very small, however, in comparison with numbers retrieved for some of the other applications, such as laparoscopic cholecystectomy. This may imply that interest is restricted to a smaller user group, or it may be just the early stage of a rising trend.

The first reference on teledermatology only appeared in Medline in 1992 *(Figure 14)* and by 1995 the number of references (4) was still very low. This suggests that teledermatology is a much more recent development than teleradiology and that either the current level of teledermatology relative to teleradiology is still quite low or that any increase in the level of teledermatology activities has yet to be reflected in the clinical literature. Another possible explanation is that when the main impact of applications is at the primary-care level, publication may be sparser relative to interest levels. Pressures to undertake research and to publish are higher in teaching hospitals than in general practice.

Discussion

Commercial organisations as drivers

A number of factors, common to telemedicine in general, appear to have influenced the timing of trials of teleradiology and teledermatology and the way in which they have been undertaken. Telemedicine has historically been technology-driven rather than needs-based and many projects have been funded by commercial technology providers. Trials have tended to focus on the technical feasibility of the system under investigation, rather than focusing on aspects such as clinical efficacy and cost-effectiveness. When results did not match physician expectations, interest waned until the technology improved, following which the cycle was repeated. To this extent the timing of trials has largely been dictated by each new wave of technological development. McLaren and Ball (1995) noted that there was much current interest in teleradiology and a drive to develop systems that would transmit digitised images of radiographs and slides at higher and higher rates for remote examination. While there were obvious attractions for radiologists in systems that allowed remote manipulation of real-time images, they maintained that the benefits to the patient and the impact on delivery of health-care were less clear. It would be interesting in a future study to determine whether other instrument-based medical technologies show a similar pattern of trials, since they are likely to have the same commercial imperatives and incorporation of generic technologies. In a number of cases, the introduction and evaluation of telemedicine seemed to be motivated, at least in part, by a desire to imitate developments elsewhere, although it may also have been driven by perceived need.

User population

Reports of telemedicine trials have tended to be anecdotal and descriptive. This may be at least partly because funding for projects is generally provided only for a year or so, and such a limited timescale would constrain the development and conduct of a comprehensive RCT. This problem is compounded by the fact that because many telemedicine projects serve sparsely populated areas, numbers of patients within the trials are low, making it difficult to generate sufficient patient contacts to arrive at statistically valid results. According to Smits and Baum (1995), only about 2000 telemedicine consultations in total were undertaken in each of the years 1993 and 1994. Perednia and Brown (1995) expected that future telemedicine studies would require collaborative efforts linking many remote sites. Hence the practical difficulties in setting up trials of a technology whose advantage is its suitability for sparse populations mitigate against evaluation. This problem is compounded when the technology is fast-changing because accumulation of data over time is compromised by the changing technical specifications of the hardware and infrastructure.

Legal issues have also influenced the development of telemedicine, especially in the USA where most telemedicine activity has taken place. In the event of damaging litigation, a reduction in funding of projects might ensue, restricting future telemedicine development. In addition, physicians in the USA are licensed state by state, which raises questions over their professional status if remote diagnosis of a patient in another state is undertaken. The question of the professional status of telemedicine consultants also exists where programmes serve different countries.

Reimbursement issues

In the USA, the Health Care Financing Administration does not provide reimbursement on a national level for telemedicine, which restricts its wider expansion. Only teleradiology and tele-pathology qualify for reimbursement by Medicare, the federal healthcare system. The Administration has recently begun a 3-year trial of Medicare reimbursement for telemedicine consultations in the states of Georgia, Iowa, North Carolina and West Virginia. It is expected to take about 5 years before the trial recommendations are assessed and transformed into legislation (*The Economist*, 11 January 1997). According to Smits and Baum (1995), the Health Care Financing Administration did foresee accessible health-care being provided through telemedicine but stated that solid data was required to ensure that quality was not

compromised. This is a chicken and egg situation since substantial data is unlikely to be available until reimbursement leads to wider use.

The current resurgence of interest in telemedicine is caused by a number of factors. The technology has improved and become more affordable. Health-care reform in many developed countries has led to an emphasis on the provision of efficient services and a demand by patients and physicians in rural areas for access to high-quality medical care irrespective of location. Widespread expansion of telemedicine, however, should be preceded by adequate evaluations of efficacy and cost-effectiveness. The outcome of such evaluations, if favourable, might encourage a more general acceptance of telemedicine by patients and physicians alike.

Chapter 7

Systematic review of diffusion and evaluation of diagnosis of genetic susceptibility to breast cancer, focusing on factors influencing timing

Summary

The purpose of this systematic review is to investigate when and how evaluations of diagnosis of genetic susceptibility to breast cancer have been undertaken and to clarify the factors that have influenced the timing of those evaluations. Literature searches, primarily on Medline and Embase, were supplemented by articles obtained by other means, such as contact with experts in the field, hand-searching of reference lists of key papers and monitoring the contents of a small number of journals in 1996.

There are about 25,000 new cases of breast cancer in the UK each year, of which up to 10% have a heritable basis. Mutations of the breast-cancer genes, BRCA1 and BRCA2, may account for at least 80% of families with a history of breast cancer. Genetic tests can determine whether an individual is predisposed to develop a disease such as breast cancer.

Genetic testing has raised a number of ethical, legal and social concerns. In the event of a positive result, there is no clear action to take which would effectively prevent the disease from developing. A negative test result does not guarantee that breast cancer will not occur. Notification of test results, both positive and negative, can cause psychological distress. In addition, individuals who opt for genetic testing may leave themselves open to insurance and employment discrimination. While it was being argued that genetic testing should remain in the research setting until these issues were resolved, companies such as Myriad Genetics in the USA were developing commercial diagnostic testing services to identify mutations of BRCA1 and BRCA2.

Genetic diagnosis for predisposition to breast cancer still seems to be at a very early stage as far as clinical trials are concerned. In 1994 in the USA, the National Centre for Human Genome Research, together with the National Cancer Institute, the National Institute of Nursing Research and the National Institute of Mental Health, jointly awarded more than $2.5 million to research groups to answer some of the questions surrounding genetic testing. Over a 3-year period, 11 research projects were to be funded. A major information-gathering and evaluation project, funded by the European Union, integrating several European centres and coordinated at the University of Aberdeen by Professor N Haites, is in its early stages. Its aims include documentation and evaluation of clinical services and management of familial breast cancer and preparation of guidelines.

Little information was found concerning timing of evaluations.

Background

Of the 25,000 patients in the UK who develop breast cancer each year, the disease has a heritable basis in up to 10%, with the most obvious indicators being early age of onset and a family history of the disease. Genetic tests can, in some cases, indicate whether a person has a predisposition to developing such a disease. Individuals might wish to be tested if there is a family history of breast cancer and they are worried about contracting the disease or passing it on to their children.

The predictive quality of the results of genetic testing depend on complex laboratory procedures and on the accurate interpretation of results. Tests can vary in their ability to detect abnormal genes, interpreting their results is often complex and they may produce false-positive or false-negative results. False-positive results indicate wrongly that the mutant form of gene had been detected, while false-negative results occur if the test failed to detect a gene mutation when it was present.

The recent isolation of genes that increase a woman's likelihood of developing breast cancer brought with it the possibility of testing large numbers of people to see whether they carried

the predisposing genes (Lerman & Croyle, 1994). Offit (1996) considered that inherited mutations of the breast-cancer gene, BRCA1, might account for between a third and a half of hereditary breast cancers and that, for women with inherited mutations of BRCA1, the lifetime risk of breast cancer might be as high as 90%. Mutations in the second breast-cancer susceptibility gene, BRCA2, conferred a similar risk of breast cancer to BRCA1 mutations but with a lower risk of ovarian cancer (Yates, 1996). He argued that mutations of the genes, BRCA1 and BRCA2, accounted for at least 80% of families with a history of breast cancer.

A number of ethical, legal and social issues surround genetic testing. Being told of genetic susceptibility to a disease such as breast cancer can affect individuals and families on a personal level, as well as affecting employment prospects and acceptability for insurance purposes. One of the main reasons why genetic testing remains controversial is that there is no clear way forward once the test results are known. It could be argued, for example, that in the event of a positive test little could be done to prevent the onset of the disease, other than prophylactic mastectomy or increased mammogram surveillance. A positive result, however, could prove psychologically very damaging for the individual concerned. One of the dilemmas in testing for BRCA1 is that the test's usefulness seems to be confined to those with a family history of the disease. The BRCA1 gene is large and, as Yates (1996) noted, more than 100 mutations had already been identified, scattered throughout the gene. Because the gene had so many mutations, testing might not be feasible for the general population.

Development and diffusion

In 1990, Marie Clare King mapped, by segregation analysis, the first breast-cancer susceptibility gene (BRCA1) to chromosome 17q21. Eeles (1996) stated that the breast-cancer gene BRCA1 probably accounted for about 2% of all cases of breast cancer, rising to 8% of all breast cancers in women under 30 years of age. A second breast-cancer susceptibility gene, BRCA2, has since been mapped to chromosome 13q12. Whereas in 1990 genetic researchers knew of only a few mutated genes that caused disease, new gene discoveries are currently being made on a weekly basis (Brown, 1996).

According to Sidebottom (1995), although the sequencing of the BRCA1 gene provided the basis for a DNA test of susceptibility for high-risk women, there was almost universal agreement that this would be premature outside a strictly controlled

research programme. There was a need to agree a coherent screening policy which might involve regulation or licensing of testing. If this was not done then market forces would predominate and it would become very difficult to introduce a rational, controlled and cost-effective programme.

Objectives

The objectives of this review were to identify the potential benefits and disadvantages of diagnosis of genetic susceptibility to breast cancer, to find out when and how evaluations of this diagnostic approach had been undertaken, to assess the extent of commercial involvement and the role of media reports, and to clarify, if possible, the factors that influenced the timing of the evaluations.

Search strategy

Literature searches were carried out primarily on the databases, Medline and Embase. The database searches were supplemented by articles identified by other means. These included contact with experts in the field, hand-searching of reference lists of key articles obtained, and monitoring the contents of a small number of medical journals in 1996 (see Appendix 1 for a more detailed description of the search strategy, criteria for inclusion and methods).

Studies included

Potential benefits of genetic testing for breast cancer

The arguments for and against genetic testing for breast cancer are complicated by the rapidly evolving body of knowledge about the genes and their relationship with the disease. In addition, psychological effects on patients are recognised as important and, since these vary between individuals and between cultural groups, they are hard to assess.

According to Lerman and Croyle (1994), one of the greatest potential benefits of genetic testing for breast-cancer susceptibility was the identification of younger women who might benefit from the initiation of mammography surveillance at an earlier age and/or on a more frequent basis. They also pointed out that notification of genetic breast-cancer susceptibility can also motivate women to adhere to recommended guidelines for breast-cancer screening.

Christensen (1996) also argued that the benefits of genetic testing for BRCA1 included the possibility of early detection through increased screening surveillance. There was also the possibility of prevention through prophylactic mastectomy or ovary removal, although the effectiveness of these approaches remained unproven. In addition, the ability to keep track of patients who tested positive could mean that they might benefit from any effective preventive measures that were developed in the future. Patients might also wish to have test results available to aid in personal decision-making; a final benefit of test results might be relief from worry if the test was negative.

As a practical guide for the present, any family with a history of cancer should be referred to a specialist family cancer clinic, such as were now available in most regions of the UK (Ponder, 1995). The clinic would confirm the family history, provide advice about the implications for other family members, and recommend screening or preventive measures where appropriate, and discuss the possibility of genetic testing.

Potential disadvantages of genetic testing for breast cancer

It is possible that at least four, and possibly more, genes predispose women to breast cancer in different families (Ponder, 1995). If no abnormality is found in the genes that had been isolated, such as BRCA1, predisposition might be due to another gene and the test would then have provided no benefit. There might also be strong and weak mutations, associated with different degrees of risk and perhaps with different cancers. In families with no extensive cancer history, the potential for mistaken prediction of risk might be considerable.

Some of the potential psychological, social and ethical issues related to screening were noted by Bryant (1996). Particular difficulties with BRCA1 screening arose from the potentially uninformative nature of the test, as many women with a negative test would still develop breast cancer, while some with a positive test would not. In addition, the particular defect would be responsible for the increased incidence in only a proportion of high-risk families. Bryant (1996) also pointed out that in the event of testing positive, mammographic screening was the only one of three commonly recommended early-detection strategies (the other two being clinical breast examination and breast self-examination) found to be effective in RCTs. Even then there was agreement on its benefit only for women aged over 50 years of age.

Lerman and Croyle (1994) argued that discussing the limitations and potential risks of genetic testing for breast-cancer susceptibility might dissuade some individuals from taking part. They argued that if genetic testing went ahead without participants' awareness of these limitations, they were likely to be disappointed with the results and might be more susceptible to the adverse psychological consequences of testing.

Christensen (1996) noted that the quandary faced by patients and physicians was that, while the test might identify someone with a potentially high likelihood of developing breast cancer in her lifetime, the question remained as to what intervention, if any, might prevent onset of the disease. Other concerns included the estimated possible false-positive rate of 5–14%, making decisions about extreme interventions like prophylactic mastectomy extremely difficult.

The potential negative consequences of testing positive included depression or the anxiety of knowing one was at higher risk for breast and ovarian cancer, physical and psychological harm that could result from prophylactic mastectomy or oophorectomy, potential discrimination in employment and insurance, possible damage to personal and family relationships, and anxiety over having possibly passed on the mutated gene to family members (Christensen, 1996). There were potentially unfavourable consequences if the test was negative, including undue reassurance, as the individual would retain the overall risk of developing breast cancer applicable to all women in the population. A different BRCA1 mutation might be carried, leading to a higher risk. There was also the possibility that a negative test result might cause the patient to feel anxiety if other family members tested positive.

The above arguments illustrate the difficulties of making straightforward decisions about the true medical potential of the technology.

Commercial involvement

Eeles (1996) stated that although predictive genetic testing for BRCA1 would remain in the domain of the cancer geneticist for the next few years, it would soon impact on general oncological practice. It was therefore important that the issues of insurance for carriers of altered breast-cancer genes and information for women at risk should be addressed. It was important that the implications of the test were understood and it was also extremely important that genetic tests were not sold direct to the public; test results should be conveyed only by those who understood their implications.

Eeles' prediction has already been overtaken by events. Even at this early stage in the development of genetic testing, business organisations have already signalled their intention to provide such tests on a commercial basis. Because of the problems discussed earlier, this has caused considerable controversy.

McCarthy (1996) reported that the Genetics and IVF Institute, a US company, had begun to offer screening for a mutation in the breast-cancer gene, BRCA1, despite the fact that many leading genetic researchers felt that testing for the mutation should be used only in the research setting until the true risk of having the mutation was determined.

Christensen (1996) noted that the US company Kaiser Permanente were developing an evidence-based clinical practice guideline for the testing of individuals who may be carriers of a gene mutation, BRCA1, which is linked to the development of breast and ovarian cancer. Part of the guideline project was to include a confidential BRCA1 registry to track Kaiser Permanente members who decide to undergo BRCA1 testing or who may be candidates for the test in the future. Kaiser Permanente claimed that the guideline would also address the need for psychological and social support for each patient.

Myriad Genetics, a genetic testing company in Salt Lake City, Utah, USA, recently announced it would begin offering large-scale testing in late 1996 to detect a mutated gene for inherited breast cancer, BRCA1. It also announced that it was intending to develop a test that incorporated the BRCA2 gene, thus enhancing the ability to identify genetic predisposition to hereditary breast cancer. *Health Service Journal* (14 November 1996) reported that Myriad Genetics was putting a genetic testing kit on sale which would tell a woman whether she was at a high risk of developing breast or ovarian cancer. The test, which would use blood samples to look for mutations in the tumour suppressing genes, BRCA1 and BRCA2, was to cost $2400 a time. The report pointed out that family doctors would lack the counselling expertise to advise patients who had tested positive, and what medical action was to be taken on testing positive was far from clear. Another company, OncorMed, was already offering limited BRCA1 testing services for use in research settings for women at high risk.

Although companies such as OncorMed and Myriad either are currently offering or plan to offer testing for BRCA1, the usefulness such a test for the general population remains controversial. The two forms of genetic predisposition that have been discovered (BRCA1 and BRCA2), account for less than 10% of all breast cancers. For a woman without a family history of breast cancer, negative results are by no means a guarantee that she will not develop breast cancer. In addition, some women who test positive for BRCA1 or BRCA2 will not go on to develop breast cancer.

Another issue is the cost of the test and who should be responsible for payment. Some insurance companies might pay for such tests and then raise premiums for those who tested positive. Some people might therefore prefer to pay for the tests themselves to avoid the risk of having their premiums raised or their policies cancelled in the event of their insurance company becoming aware of their test results.

Regulation

In the USA, most new genetic tests do not need Food and Drug Administration (FDA) approval. If a biotechnology company researches and uses a genetic predisposition test within its laboratories, then FDA approval is not required. As an alternative to FDA approval, companies such as Myriad Genetics are establishing institutional review boards to guide clinical protocols. Josefson (1996) noted that the FDA has stated that it has the authority to regulate genetic testing but currently lacked the staff to do so.

Evaluation of genetic testing for breast cancer

Little evaluation has yet been initiated but considerable thought has been given to ways of approaching the problems of genetic testing in general, and some writers have considered the breast-cancer case in particular.

Recommendations

Eyre and colleagues (1995) argued that the specific mutations of BRCA1 that increased susceptibility should be defined at the molecular level by studying the blood and tumour samples from those individuals already under study in large cooperative groups and in the International Familial Breast Cancer Consortium. Once the mutations which accounted for a large percentage of the total number of expected breast cancers in this group were identified, it would be feasible to expand the testing to the high-risk families already identified in the process of searching for BRCA1. They recommended that studies should focus on positive predictive value, accuracy, reliability, sensitivity and cost. They also suggested that before widespread screening occurred, a limited cohort of people in

the general population with and without a family history of breast cancer should be studied to determine whether the BRCA1 mutations that occurred in families with a history of cancer actually carried the same risk of breast cancer as the general population. It would also be necessary to address ethical and practical issues related to false-positive results, true-positive results, susceptibility to radiation and whether it was possible to influence the course of the disease (Eyre *et al,* 1995).

A conceptual model for genetic testing for breast-cancer susceptibility was described by Lerman and Croyle (1994). This model addressed the three key psychological and behavioural issues of ensuring informed consent for testing, minimising adverse psychological consequences, and promoting breast-cancer prevention and screening practices.

As genetic testing for breast cancer was initiated on a larger scale, it would be critical to conduct carefully designed studies to evaluate the impact of these programmes on psychological status and cancer prevention and control practices. Lerman and Croyle (1994) maintained that prospective studies were essential to determine the psychological and behavioural impact of genetic risk information. Assessments should be conducted at multiple time-points, including before testing, immediately after notification, at short-term follow-up (3 months), and at long-term follow-up (1 year). Critical outcome variables should include psychological status, marital and family functioning, quality of life, health behaviours, reproductive intentions, and healthcare use. Lerman and Croyle noted that much of the early work on genetic testing and counselling had focused on participants in early pilot programmes, a group that may be highly self-selected and whose responses might not be representative of the entire at-risk population.

The second stage of psychological research on genetic testing for breast-cancer susceptibility should include controlled clinical trials of different counselling protocols (Lerman & Croyle, 1994). Because of unique aspects of breast-cancer causes and prevention, counselling protocols developed for traditional hereditary diseases might have limited applicability to genetic testing for breast cancer. As a result, it would be important to develop and test new counselling approaches that varied in terms of the content, process and timing of delivery of genetic information.

Stix (1996) reported that, in the USA, the National Breast Cancer Coalition, the American Society of Human Genetics and the National Advisory Council for Human Genome Research had recommended that testing be conducted only as part of an on-going research effort until issues such as potential discrimination by health insurers and employers could be resolved. Nevertheless, it was expected that the tests developed by Myriad Genetics and OncorMed would come into clinical use in a few years' time. Professional groups, such as the American Society of Clinical Oncology, were already breaking with the medical consensus to oppose clinical testing outside a research study by recommending that testing be permitted for anyone with a family history of breast cancer. Stix also reported that the National Cancer Institute had recently established a National Cancer Genetics Network as a means by which patients could join a research study being undertaken, and receive genetic testing and counselling.

Evaluations of attitudes

Watson and colleagues (1995) reported the early experience of their group in BRCA1 testing by linkage. The eligibility criteria were that there should be at least four cases of breast and ovarian cancer, including at least one ovarian cancer and two young-onset breast cancers (below 50 years of age) within a family. Two families (32 unaffected individuals; 17 female, 15 male) entered the programme and 25 individuals expressed an interest in taking part, but eight did not proceed. Of the original 25, 11 women and four men attended for pre-test counselling, two of whom did not proceed to testing, leaving 13 individuals (ten female, three male) who were tested. The main reasons for wanting the test were to help research and clarify the situation for their offspring. Watson and colleagues (1995) reported that general mental health assessment indicated a slight increase in morbidity at about the time of blood sampling. All participants, however, reported that the information on genetic testing was helpful. It was recommended that individuals with an unfavourable result should be offered psychological support, and that genetic testing for a breast/ovarian cancer gene should be offered within a multidisciplinary team.

A study that examined interest in and expectations about the impact of a potential genetic test was reported by Lerman and colleagues (1994). The participants were 121 first-degree female relatives of ovarian cancer patients. The study design was cross-sectional. Participants completed a structured telephone interview on attitudes to cancer and genetic testing, and self-reporting psychological questionnaires to assess coping style and mood disturbance. Overall, 75% of first-degree relatives claimed that they would definitely want to be

tested for BRCA1 and 20% that they probably would. One limitation of the study was that the outcome was intention to receive a hypothetical genetic test, rather than to participate in an actual test. In addition, the sample size was small and, as participants were predominantly white and middle class, the results might not be generalisable to high-risk women in the wider community (Lerman *et al,* 1994).

The conclusions of Lerman and colleagues (1994) were that the demand for genetic testing among first-degree relatives of cancer patients was likely to be great, and that those who elected to participate might represent a more psychologically-vulnerable subgroup of high-risk women. They recommended that before genetic testing was integrated into routine clinical practice, effective and ethical means of obtaining consent and communicating genetic risk information should be identified. In addition, the development of clinical protocols should be guided by empirical research that examined the psychological impact of testing on participants and their families. Consensus guidelines for surveillance and prevention of breast cancer among carriers also needed to be established, including validated methods for enhancing patient adherence. They concluded that these issues would be best addressed if genetic testing for the BRCA1 gene initially was conducted within the context of research that carefully assessed the immediate and long-term impact of risk notification (Lerman *et al,* 1994).

Evaluations under way
An article in the journal *Oncology* (volume 8, 1994) reported that, in the USA, the National Centre for Human Genome Research, along with the National Cancer Institute, the National Institute of Nursing Research, and the National Institute of Mental Health, had jointly awarded more than $2.5 million to research groups to help answer some of the questions surrounding genetic testing. The 3-year grants were to support 11 research projects in a consortium coordinated by the National Centre for Human Genome Research's Ethical, Legal, and Social Implications Branch. The consortium format was designed to allow investigators to compare findings along the way on issues common to all of the projects, to reduce duplication of effort in the research, and to promote sharing of information.

The Fred Hutchinson Cancer Research Centre, Seattle, was to provide genetic counselling and DNA testing for the BRCA1 mutation in women from families at increased risk for breast cancer. This group would examine the impact of

alternative forms of counselling on womens' perception of breast-cancer risk, decision-making about DNA testing, and fears about breast cancer. In a second project, the University of Washington in Seattle would gather information about genetic testing for cancer risks from women receiving 'routine' health-care, genetic service providers and primary care providers; the University team would study how women reacted to receiving risk information on breast cancer based on their family history; and whether primary-care providers differed from genetics professionals in their approach to counselling about the genetic risks for breast cancer.

The Fox Chase Cancer Centre, Philadelphia, was to study the use of oncology nurses as the primary source of patient counselling about DNA testing for BRCA1 in an ethnically diverse population. The University of Utah, Salt Lake City, would study how adolescent girls were affected by their parents' testing for the BRCA1 gene. The Dana-Farber Cancer Institute, Boston, Massachusetts, was to develop, implement, and evaluate a programme in which teams of genetic counsellors and nurses were educated to administer education and counselling to families with the BRCA1 gene.

The University of Hawaii, Honolulu, was to study an array of factors that motivated or inhibited multi-ethnic Hawaiin residents to seek DNA testing for colon-cancer susceptibility; they would also look at factors that affected intentions among multi-ethnic health professionals to offer DNA testing for cancer risk. The Johns Hopkins University, Baltimore, was to assess both healthcare provider and patient expectations of what information would be discussed when obtaining or giving informed consent for predictive cancer testing. Based on their findings, the group would then develop a model informed consent protocol for use in BRCA1 testing. This group would also to study social and psychological factors behind a decision to be genetically tested in people with a family history of colon cancer, and would evaluate the impact of test results on cancer surveillance and prevention behaviour.

The Anderson Cancer Centre, Houston, Texas, was to characterise the psychosocial and behavioural impact of DNA testing for hereditary nonpolyposis colon cancer. Georgetown University, Washington DC, was to study methods for educating and counselling women with a family history of breast or ovarian cancer who seek DNA testing for cancer predisposition, and to assess the impact of alternative education and counselling strategies on knowledge, decisions, psychological well-being,

and health behaviour. Finally, the Princess Margaret Hospital, Toronto, Canada, was to focus on the healthcare providers who delivered information to patients seeking DNA testing for heritable cancers. The results of this study would form the basis for guidelines on disseminating health-risk information to people undergoing DNA testing.

A major information-gathering and evaluation project, funded by the European Union, integrating several European centres and coordinated at the University of Aberdeen by Professor N Haites, is in its early stages (Personal communication, 1997). Its aims include documentation and evaluation of clinical services, management of familial breast cancer and preparation of guidelines.

Economic aspects

The costs of widespread testing have evoked concern. Warshaw (1994) drew attention to the costs of genetic screening for risk of breast cancer in the USA and the question of who should be responsible for payment. In addition to the cost of the test itself, there were also costs involved in the counselling required to obtain informed consent and further counselling when the results of the test became known.

Ponder (1995) noted that BRCA1 was a very large gene, with mutations scattered all over the gene and different mutations occurring in different families. The amount of work involved in scanning the gene in each new family was likely to be considerable and, in the UK, there were no government funds available to support such analysis. In addition, the fact that individuals would react in different ways to the results of genetic tests might make the adoption of standard protocols problematic, while counselling tailored to individual needs would be more time-consuming and costly. It was improbable that research laboratories would be able to undertake more than a few small research series and, in the event of commercial laboratories becoming involved, the cost to individuals could be hundreds of pounds.

The vast majority of work in the cancer predisposition field, including molecular genetic analysis and cancer genetics clinics, was research funded (Evans, 1995). The author foresaw no immediate prospect in the UK of a major government cash injection to cope with what he expected to be an inevitable increase in demand.

Until clearly defined benefits have been identified, it will be difficult to justify the costs of routinisation of this procedure.

Media impact

Genetic testing has attracted considerable media attention over the last few years. A number of significant issues have been raised. Ethical, psychological and insurance issues have been debated extensively. Testing for breast-cancer susceptibility has also received some specific coverage.

A report in the *Financial Times* (3 January 1995) noted that the potential for discrimination based on gene testing was especially strong in the USA, quoting the example of a former school bus driver who was refused health insurance four times. The report concluded that until issues such as these were addressed, the growth of gene testing in the USA would be held back. The *Financial Times* (24 January 1995) discussed genetic testing on volunteers whose families had a history of cancer in order to identify whether those individuals were predisposed to developing the disease. The report also drew attention to the attitudes of health- and life-insurance companies, and noted that the threat of 'genetic discrimination' would be a huge disincentive to population screening.

Beardsley (1996), writing in *Scientific American,* described the human genome project and developments in genetic testing, noting that genetic testing could cause psychological problems and leave an individual open to discrimination. The article drew attention to the fact that the significance of a positive test was less than clear. Nevertheless, companies such as Myriad Genetics were planning to offer BRCA1 testing to all women diagnosed with breast or ovarian cancer and to their close relatives. Beardsley pointed out that worries over health-insurance discrimination have led to a trend towards secrecy in genetic testing, or to individuals avoiding taking genetic tests altogether.

Leonard, however, writing in *Scotland on Sunday* (21 January 1996), noted concern that potential demand for genetic tests might outstrip supply, as there were only about 200 trained geneticists and support workers in Britain. She also drew attention to the potentially huge cost of providing such screening as part of the NHS, and to the dangers of health and employment discrimination as a result of such tests.

Publication trends

The annual numbers of references to genetic screening for predisposition to breast cancer identified through a search of the Medline database are shown in *Figure 15* (the MeSH term which covers genetic testing is 'genetic screening'). The search strategy retrieved one early reference to the subject
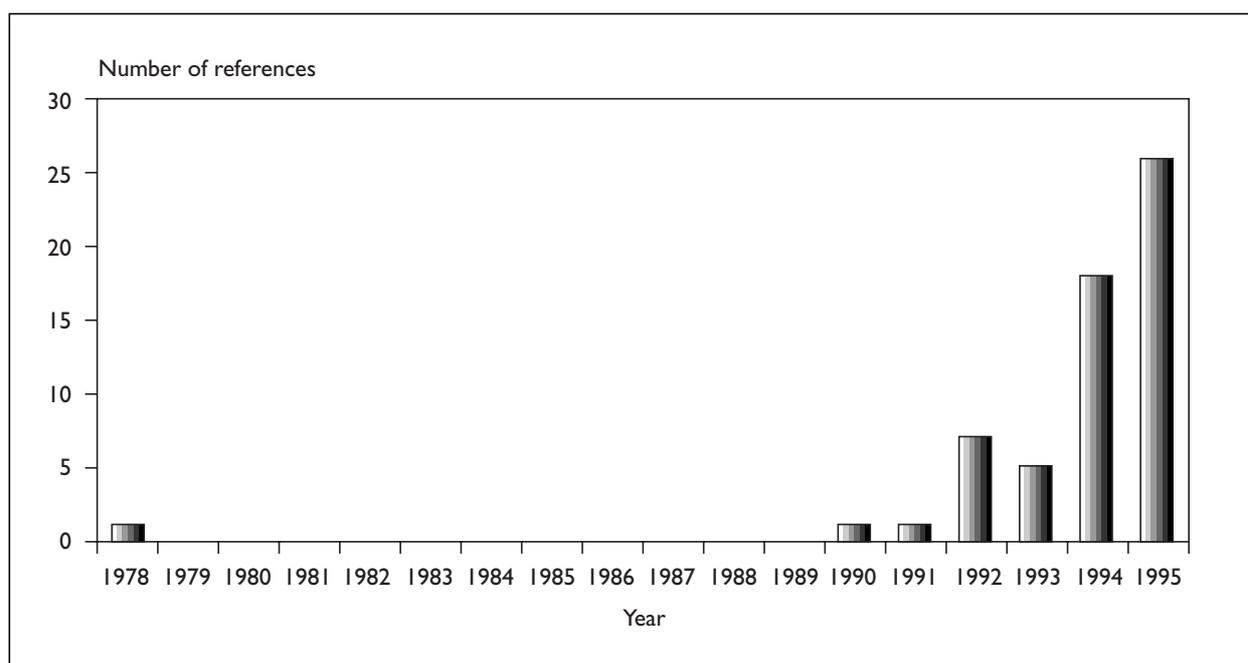
FIGURE 15 *Genetic testing for breast cancer – Medline references*

in 1978 (Purtillo *et al*, 1978), in which the basic genetic mechanisms responsible for tumour formation were discussed. Other than this, references first began to appear in 1990, with the annual numbers exhibiting a generally upward trend. Total annual numbers are still relatively low, however, with the figure for 1995 standing at 26 (as at October 1996). This suggests that the development of genetic screening for predisposition to breast cancer is still at an early stage in the diffusion process.

Discussion

The development of diagnosis of genetic susceptibility to breast cancer has raised a number of important issues which evaluations need to address. The ethical issues raised are complex and require further exploration. The information obtained is prone to inaccuracies. The interpretation is difficult to communicate to the patient and, in addition, no definite benefits can be offered. As Christensen (1996) argued, a thorough process of informed consent is an essential element of any genetic testing procedure. This means that potential users need to receive and understand information not only about the potential benefits but also the limitations and potential risks of the technique. This is not easy to achieve in this case.

Guidelines should be developed for communicating genetic information about breast cancer susceptibility or for providing recommendations and follow-up care for identified gene carriers. As Lerman and Croyle (1994) indicated, these are needed because notification of cancer risk could have serious negative psychological consequences. Also, in the absence of proper counselling and follow-up, psychological distress might potentially undermine adherence to recommendations for surveillance and possible prevention. Special training is also required for physicians so that they can communicate information about genetic testing to affected individuals and their families.

There is also a need to protect patient privacy and the confidentiality of test results. Although confidentiality is required for all medical information, it has particular relevance for genetic testing because of the potential for discrimination in employment or for insurance purposes (Christensen, 1996).

The process of evaluation is still at an early stage. The need has been recognised and initial assessment projects have been funded. However, all those involved are struggling with the complexity and diversity of the issues they must address.

Chapter 8

Systematic review of diffusion and evaluation of gene therapy for cystic fibrosis, focusing on factors influencing timing

Summary

The purpose of this systematic review was to investigate when and how evaluations of gene therapy for cystic fibrosis have been undertaken and to clarify the factors that have influenced the timing of those evaluations.

Literature searches were undertaken, primarily on Medline and Embase. The database searches were supplemented by articles obtained by other means, such as contact with experts in the field, hand-searching reference lists of key papers, and monitoring the contents of a small number of journals for 1996.

Cystic fibrosis is the most common lethal hereditary disorder among Caucasians, occurring once in approximately 2500 births. It is caused by a chromosomal abnormality which affects chloride transport across epithelial membranes, resulting in abnormally thick mucus. Individuals who have cystic fibrosis tend to suffer from chronic lung infections and lung damage is the usual cause of premature death, with few patients surviving beyond 30 years of age. Gene therapy is a method of treating diseases by replacing the defective gene with a copy of a normal gene.

Evaluations of gene therapy for cystic fibrosis include those by Zabner and colleagues (1993), Crystal and colleagues (1994), Knowles and colleagues (1995), and Caplen and colleagues (1995). These were Phase I studies with small numbers of patients, in which adenovirus-mediated vectors were used apart from the study by Caplen and colleagues, who used liposome-mediated vectors. A difficulty with the use of an adenovirus was the possibility of inflammation of the lower respiratory tract. The studies generally reported a temporary improvement in function which then returned to pretreatment levels.

By mid-1996 there were five clinical protocols in progress for gene therapy for cystic fibrosis in Europe, and 11 in the USA. Adenovirus-mediated gene transfer has tended to be used for initial trials in the USA, while liposome-mediated gene transfer has tended to be more popular in the UK and the rest of Europe.

The reports of trials provided little or no information about when they took place or the reasons for their timing. Interview data and published commentaries indicated that factors which impacted on the timing of these trials, however, included the strict and still-evolving centralised regulatory process through which they had to pass, official concerns to ensure that a major new technology be carried through to application, high levels of popular interest and concern about genetic technology in medicine, the fact that gene therapy is still at an early stage of development, the costs involved, and technical obstacles to producing enough genetic material for transfer.

Background

Cystic fibrosis is a single gene disorder characterised by abnormal salt and water transport that leads to abnormal airway secretions, impaired mucociliary clearance, chronic bacterial infection, and premature death. A variety of epithelial tissues are affected in this disease, including airway, pancreatic, sweat ductal, and gastrointestinal epithelia. However, lung disease is the major cause of morbidity and mortality in this disorder. For this reason, and because of the relative ease of access to lung tissue, initial gene therapy efforts have been directed towards lung disease (Johnson, 1995).

Because cystic fibrosis is an autosomal recessive disorder, the introduction of a normal copy of the gene into the host cell should result in normal transport function (Johnson, 1996). Retrovirus-mediated gene transfer is the method used in most currently approved human gene transfer trials in the USA, although liposome-mediated gene transfer is more popular in European studies (Johnson, 1996). The factors affecting vector choice are discussed below.

Wivel (1994) described the various methods of gene transfer, noting that the demonstration that retroviruses could transduce the vast majority of dividing cells in a cell-culture system marked an important step forward. Retroviruses can be integrated stably into the genome of the host cell and do not cause cell death as a result of infection. To use this virus system for gene transfer, the structural genes of the wild-type retrovirus are removed and replaced with the therapeutic gene. Through the use of specially designed packaging cell lines, retroviral particles carrying the therapeutic gene are produced, but these particles are unable to replicate properly. Thus, the target cells become infected and the vector carrying the therapeutic gene is stably integrated but no infectious virus is produced. Retroviruses can only infect dividing cells, thus creating the problem of insertional mutagenesis. For example, the integration site could be involved in disrupting the function of a normal gene, activating a proto-oncogene, or inactivating a tumour-suppressor gene. The latter two events could predispose the patient to the development of a malignancy (Wivel, 1994).

Another viral vector that has been approved for use in patients is the adenovirus. Recombinant adenoviruses will infect non-dividing cells, and they can be prepared to significantly higher concentrations than retroviruses. Adenovirus vaccines have been used in humans and have been shown to be safe in this setting. Constructs of an adenovirus and the cystic fibrosis transmembrane conductance regulator (CFTR) gene have been created for delivery into the lungs of patients with cystic fibrosis (Wivel, 1994).

As an alternative to viral vectors, several physical methods can be used to deliver DNA to cells. Such approaches avoid some of the problems associated with the use of infectious agents such as viruses, for instance, immunogenicity, but it remains to be determined whether these approaches are sufficiently efficient and can produce long-term expression of the desired gene product. One such method for direct *in vivo* gene transfer uses DNA that is complexed with cationic lipids or liposomes. Proposed gene therapy studies on patients with cystic fibrosis, to be performed in the UK, will use a liposome-DNA mixture to deliver the CFTR gene to the epithelial cells of the lungs (Wivel, 1994).

Development and diffusion
Hillman (1996) reported that the first clinical trial of gene therapy began in 1990, and that since then, more than 100 protocols involving gene transfer have been approved by the US National Institutes

of Health. The majority of clinical trials currently approved and on-going in the USA involve gene therapy in the treatment of various cancers.

As of mid-1996, there were five protocols for gene therapy for cystic fibrosis studies in Europe and 11 in the USA. Most of the US studies used an adenovirus, while a liposome-mediated vector was the method of preference in Europe.

Wivel (1994) stated that a number of changes were needed before human gene therapy could have widespread clinical use in the treatment of a disease. Gene therapy currently is a high-technology treatment that is limited to a few medical centres. It is labour intensive, requiring a molecular biology laboratory and employees with expertise in virology to prepare vectors.

Objectives

The objectives of the review were to identify when and how evaluations of gene therapy for cystic fibrosis had been undertaken and to clarify the factors which influenced the timing of those evaluations, including regulatory issues, commercial involvement and media impact.

Search strategy

Searches for papers on gene therapy for cystic fibrosis were made primarily on the databases Medline and Embase. The database searches were supplemented by papers identified through other means. These included contact with experts in the field, hand-searching reference lists of key articles, and monitoring the contents of a small number of medical journals, including *Human Gene Therapy,* for 1996 (see Appendix 1 for a more detailed description of the search strategy, criteria for inclusion, and methods).

Studies included

Development of regulatory processes
Regulation of gene therapy is still evolving. While regulators in the USA and Europe are imposing a similar regime to that applied to ethical pharmaceuticals, special attention is being given to a close scrutiny of ethical issues, and to considering whether new issues are raised by the nature of invasive use of genetic technologies. Wivel (1994) commented on some of the ethical issues involved in gene therapy. Somatic-cell gene therapy is the

attempted correction of genetic defects in any cells of the body except germ or reproductive cells. The ethical issues surrounding germ-line intervention are particularly complex. From a purely practical viewpoint, germ-line therapy might be more efficient than somatic-cell gene therapy because it allows correction of the genetic defect in the patient and in future generations, while somatic-cell intervention must be repeated for each generation. However, risks in the somatic-cell approach are limited to one patient, whereas genetic 'mistakes' in the germ-line approach might be propagated in subsequent generations. There is also the unsettling possibility of using germ-line technology for enhancement as well as treatment of disease by, for example, using genetic intervention to improve physical capabilities, intelligence or physical appearance (Wivel, 1994).

The Medicines Control Agency, formerly the Committee on the Safety of Medicines is the UK statutory body responsible for all aspects of medicinal products for human use. Assessment of protocols by this agency is currently a prerequisite of all gene-therapy trials in the UK.

The ethical issues underlying the principles governing regulation of gene therapy trials and humans were described by a UK Government Committee (the Clothier Committee) which was formed in 1989 and reported in 1992 (Clothier, 1992). This group continued to act as a Genetic Therapy Advisory Committee until the establishment of the Gene Therapy Advisory Committee in 1993 (Caplen *et al*, 1994). This Committee is seen as advising on both the ethics and the scientific merits of gene therapy proposals. As well as the Gene Therapy Advisory Committee, local hospital ethics committees will continue to consider the overall plan of the programmes of work and the conduct of trials, with particular reference to safety, discomfort, informed consent, long-term follow-up and the protection of the confidentiality of those taking part (Caplen *et al*, 1994).

In the USA, the Recombinant DNA Advisory Committee (RAC) was established to develop guidelines that establish the safest conditions for different types of genetic manipulation and to provide oversight of the application of these guidelines (Zallen, 1996). The RAC was set up in 1974 as a committee of scientists, physicians, ethicists, lawyers and consumer representatives (Ross *et al*, 1996). There have been recent calls for its dissolution on the grounds that it was an anachronism whose existence was hampering progress in the field,

and that there were other mechanisms in place to review gene therapy experiments, such as local Institutional Review Boards, and the FDA (Zallen, 1996). However, Zallen (1996) commented, the serious inadequacy of many protocols submitted to the RAC showed that Institutional Review Boards were not carrying out their function adequately; many of the informed consent documents submitted with protocols were badly written and structured, contained confusing, highly technical jargon which might mislead participants into thinking an experiment was a cure, and were often vague as to what risks were involved.

Commercial involvement

Concerns have been expressed that early entry of commercial organisations into gene therapy development might have undesirable effects. Panellists involved in a roundtable discussion of socio-economic and ethical issues of gene therapy concluded that this field of study needed time to mature scientifically without pressure to develop a marketable therapeutic product (Hillman, 1996). Another factor to be considered was that gene therapy also held threats as well as promises for the traditional pharmaceutical research and development market. If gene therapy effected a cure for a specific disease, this would inhibit the market for drugs for that disease (Hillman, 1996).

These concerns have not prevented companies targeting this area for product development. A number of large pharmaceutical companies have made investments in this technology, including SmithKline Beecham and Novartis (for further information, see the annual reports of these companies). Many biotechnology companies have also made major commitments. According to Wivel (1994), 17 biotechnology companies in the USA and Europe had made a commitment to the development and commercialisation of gene therapy. Although most of these companies have research commitments, some firms had chosen to focus on provision of services. Under the service arrangement, cells would be sent to the company, the gene insertion procedures carried out, the quality-control studies performed, and the cells returned to the physician for reinfusion into the patient (Wivel, 1994).

The US National Institutes of Health are spending approximately USA $135 million each year funding extramural and $60 million on intramural programmes in gene-transfer research (Touchette, 1996). Hillman (1996) noted that, as far as commercial involvement was concerned, the need for short-term return in biotechnology

financed by venture capital is at odds with the early development of the gene therapy field and the need for more fundamental research in core gene-transfer technologies. Private industry is spending on average $200 million annually on gene research (Touchette, 1996).

From the standpoint of the venture capitalists who provide funding to the biotechnology companies, gene therapy is a high-risk undertaking with no products on the market and no guarantees of commercial success. Most of the biotechnology companies are relatively small, the research is labour-intensive and the costs of the research are high. As a result, a pattern of alliance has already begun to develop in which biotechnology companies develop agreements with large pharmaceutical firms (Wivel, 1994).

The biotechnology companies' links are not just with the large drug companies; they also have many close relationships with academic laboratories. Wivel (1994) reported that, in the USA, there are a number of reasons for the close ties between the academic community that is researching gene therapy and private biotechnology companies. One stimulus for the development of these ties was a National Institutes of Health policy to promote technology transfer, which encouraged the creation of a number of collaborative research and development agreements. Under such agreements, private companies provide financing for part of the research in the government or academic institution and, in return, receive any patent rights that result from the research. Almost all of the 58 protocols approved by the RAC have some support from biotechnology companies (Wivel, 1994). Thus the links between academic researchers and commercial organisations are being fostered by US Government policy, which now strongly favours the leveraging of its own considerable investment in the technology by private sources of capital.

Factors affecting commercial success

It is not yet clear whether gene therapy will ever provide efficacious and cost-effective treatments. Apart from the uncertainties associated with the still-underdeveloped technology, there are a number of factors that will affect its ultimate commercial success.

At least three factors would strongly affect the ultimate commercial success of gene therapy (Wivel, 1994):

- the cost of making the products

- the market size (gene therapy will need to be effective for major diseases)
- the resolution of the patient issues related to gene sequences, viral vectors, cell lines, and other gene-transfer systems.

There is a requirement to produce high stocks of virally derived vectors to a suitable level for *ex vivo* or *in vivo* human gene transfer (Caplen *et al,* 1994). Large-scale production is still limited, however, by the need to use packaging cell lines and the difficulty of defining the maximum acceptable level of contaminating wild-type virus. In contrast, the plasmid DNA required for liposome-mediated gene transfer has been widely used and routinely prepared in the laboratory for many years (Caplen *et al,* 1994).

Media impact

Genetic technology in medicine has received wide coverage in the media in recent years. Gene therapy has been one of the applications that has attracted detailed interest. Touchette (1996) reported that the heightened publicity about gene therapy coming both from researchers and the media was already prompting patients into making unwise and even life-threatening treatment decisions based on false promises. Overzealous reporting of gene therapy's success when no such success yet exists might lead to the undercutting of public support for what might be highly promising success. An article in *The Economist* (16 December 1995) also drew attention to the fact that although gene therapy was promising, its achievements to date had been overstated.

Welsh and Smith (1995), writing in *Scientific American,* described how cystic fibrosis affects sufferers and gave details of gene therapy developments. Reports retrieved by the Promt search also contained details of existing treatment for cystic fibrosis along with current advances in gene therapy. The growing media coverage of cystic fibrosis, already high compared with the non-genetic applications, was noted in chapter 3. It will be interesting to observe whether this appears to impact on assessment as the technology continues to develop.

Evaluation of gene therapy for cystic fibrosis

As of June 1996, there were five protocols for gene therapy for cystic fibrosis studies in Europe, two using adenovirus and three using cationic liposomes (European Working Group on Human Gene Transfer and Therapy Central Office, 1996). Ross and colleagues (1996) noted 11 different protocols for gene therapy for cystic fibrosis studies in the USA. Almost all of the investigations have used an

adenovirus as the vector, although in one study, approved by the RAC but awaiting FDA approval, it was proposed to use a liposome vector.

The need for long-term reporting of side-effects of such a novel approach has been more readily appreciated than in the case of more established technologies. A roundtable discussion involving specialists in gene therapy (Hillman, 1996) commented that one unique aspect of gene-therapy studies would be the on-going concern about long-term adverse side-effects, since once a gene was administered, it was difficult to retrieve it.

Caplen and colleagues (1994) described the regulatory procedure used by their group to produce resources for the first stage of a Phase I trial of liposome-mediated gene therapy for cystic fibrosis. They drew attention to the fact that current laboratory studies strongly suggested that DNA transfer may, in time, have a significant impact not only on the treatment of single-gene disorders but also on other more complex diseases, and that there was increasing pressure to design clinical protocols using this technology.

Research into somatic gene therapy has advanced at an astonishing pace (Dorin, 1996). Dorin argued that preclinical trials in relevant animal models were vital for developing and refining safe and effective strategies. Clinical trials had to be undertaken at an early stage, as patients would present species- and disease-specific challenges, and data from such trials should then form the basis for further rounds of laboratory-based research.

Caplen and colleagues (1995) argued that the development of gene therapy for cystic fibrosis patients required assessments both of the safety of the gene transfer system and the effects of transgene expression on functional end-points such as the altered electrophysiology characteristics of cystic fibrosis patients. Dorin (1996) commented that most clinical trials of somatic gene therapy for inherited disease had been directed at cystic fibrosis. Two approaches had been used to correct the chloride-conducting defect in the disease: adenoviral vectors and DNA-liposome complexes. Both involved introducing the normal CFTR gene into the cells of the respiratory tract. Because the surface cells of the respiratory epithelium are terminally differentiated, this delivery strategy requires repeated treatment as new cell populations arise. Adenovirus-mediated gene transfer has tended to be used for initial clinical safety and efficacy trials in the USA, while liposome-mediated gene transfer has been the method

of choice for initial clinical safety and efficacy trials in the UK.

The target tissue for gene transfer that is most likely to be of therapeutic benefit in cystic fibrosis patients is the airway epithelium (Zabner *et al,* 1993). Because removal of airway epithelial cells, transfer of cDNA *in vitro,* and reimplantation of the cells into the lungs appear impractical, gene therapy will require treatment of airway cells *in vivo.* Zabner and colleagues reported a trial which administered an E1-deficient adenovirus, encoding CFTR, to a defined area of nasal airway epithelium in three individuals with cystic fibrosis. The authors elected to examine the efficacy and safety of an adenovirus vector in the nasal epithelia of patients with cystic fibrosis because this tissue had a morphology and function similar to those of intrapulmonary airways and because nasal epithelium manifests the cystic fibrosis chloride transport defect. A reduction in baseline potential difference (which is characteristically elevated in cystic fibrosis patients) was seen in the area of nasal mucosa exposed to the adenovirus vector. In addition, cAMP-mediated chloride secretion was stimulated compared with pretreatment values. These changes returned to pretreatment levels after about 21 days. No evidence of viral replication or virus-associated adverse effects was found, even at the highest dose tested. The major conclusion of the study was that *in vivo* application of a recombinant adenovirus encoding CFTR can correct the defect in airway epithelial chloride transport that is characteristic of cystic fibrosis epithelia.

Zabner and colleagues (1993) noted that their study contrasted with most earlier attempts at gene transfer to humans, in that they administered a recombinant viral vector directly to humans, rather than using an *in vitro* protocol involving removal of cells from the patient, transduction of the cells in culture, and reintroduction of the cells into the patient. They postulated that additional studies using nasal epithelia could be performed to answer some of the questions about adenovirus and to compare different means of delivery but, inevitably, other studies would have to use other tissues, including intrapulmonary airway epithelia in cystic fibrosis patients. No information was provided in this report about the timing of the trial.

Crystal and colleagues (1994) reported that they had administered a recombinant adenovirus vector (AdCFTR), containing the normal human CFTR cDNA, to the nasal and bronchial epithelium of four individuals with cystic fibrosis. A transient systemic and local syndrome was observed in the

bronchial epithelium, which resolved with symptomatic therapy and prophylactic antibiotics, and was most likely caused by vector-induced inflammation of the lower respiratory tract. Follow-up at 6–12 months demonstrated no long-term adverse effects. They stated that they had demonstrated that it was feasible to use an adenovirus vector to express the normal human CFTR cDNA in the epithelium of the respiratory tract of cystic fibrosis patients. Crystal and colleagues (1994) pointed out that their study did not address whether such therapy would be successful in preventing the respiratory manifestations of the disease, chronicity of expression or whether repeat administration would yield expression of the normal CFTR cDNA.

As far as the timing of the evaluation is concerned, Crystal and colleagues had proposed a clinical trial in 1992 to evaluate the administration of a CFTR cDNA adenovirus vector to the epithelium of the respiratory tract of individuals with cystic fibrosis. Their primary goals were to evaluate safety and to demonstrate *in vivo* gene transfer following administration of an adenovirus vector to the respiratory epithelium, with escalating doses to different individuals. They had begun the first human gene therapy trial for cystic fibrosis on 12 April 1993 (Crystal *et al,* 1994).

Crystal and colleagues commented that an important lesson was that, despite extensive planning, animal studies and thorough review, preclinical studies did not necessarily predict the response of humans (particularly individuals with disease) to gene therapy vectors. Despite the lack of clinically evident toxicity observed in animal studies, studies in humans, with viral vectors such as AdCFTR, should be approached with caution. Nevertheless, they maintained that only studies in humans would permit the definition of the 'efficacy–toxicity' window relevant to gene therapy (Crystal *et al,* 1994).

Knowles and colleagues (1995) performed a double-blind vehicle-controlled study to assess the efficacy and safety of gene transfer to the nasal epithelium of 12 patients. An adenoviral vector was selected for the study, which used a dose-escalation protocol. They reported no toxic effects at the lower dose of vector but at the highest dose there was mucosal inflammation in two of three patients. They found molecular evidence of low-efficiency gene transfer and expression of the normal CFTR mRNA in nasal epithelium, but there was no significant functional correction of abnormalities in ion transport. They concluded that the problem could not be overcome simply by increasing the dose of the vector, because the highest dose used was associated with inflammatory responses in two of three patients, also noting that in studies in animals a wide spectrum of toxic effects at higher doses was indicated. No information was provided about when the trial took place.

Caplen and colleagues (1995) referred to the trial undertaken by Zabner and colleagues (1993), in which one patient had shown evidence of transient adenovirus vector-induced inflammation of the lower respiratory tract, and there was also the possibility of a reduction in transgene expression as a result of developing immunity to the virus on repeated administration of the vector. Because of this, Caplen and colleagues argued that it was clearly important that alternatives to viral-based systems of gene delivery were assessed.

Caplen and colleagues (1995) reported the results of the first placebo-controlled double-blind trial assessing the safety and efficacy of liposome-mediated CFTR cDNA transfer to patients with cystic fibrosis. Nine male patients received cationic liposome and six male patients received only liposome to the nasal epithelium. No evidence of treatment-related toxicity was seen, and Caplen and colleagues argued that the absence of any clinical or histological changes correlating with the treatments administered suggested that topical application of DNA liposome complexes, at least up to the quantity administered, was safe. A partial (20%) restoration of the deficit between cystic fibrosis and non-cystic fibrosis patients was seen for the response to low chloride ion perfusion following CFTR cDNA administration. This was maximal around day 3 and had reverted to pretreatment values by day 7.

The level of CFTR transgene expression and the degree of correction of the electrophysiological defect that would be required for therapeutic benefit were unknown, according to Caplen and colleagues (1995). They concluded that further studies should aim to improve the efficiency of gene delivery and optimising gene expression. The trial took place over a period of 42 days, including recruitment (14 days) and evaluation (28 days), with long-term follow-up of patients continuing at 3-monthly intervals. No information was provided about when the trial took place.

Ross and colleagues (1996) reviewed the progress reports of gene-transfer clinical trials taking place in the USA as of June 1995. They noted that most of the work in gene therapy for single-gene

inherited disease has focused on cystic fibrosis. They reported that the approach used in the 11 protocols (by eight different investigators) and applied to 53 patients was to transfer the normal membrane channel gene, which is defective in CFTR, into cells of the respiratory tract. Almost all of the investigations had used adenoviruses as the vector although one, approved by the RAC but pending FDA approval, proposed to use a liposome vector. The situation at the present time, according to Ross and colleagues, was that in a minority of patients treated the transferred gene had been expressed in a clinically relevant location for times between 9 and 14 days, although biological and clinical improvement could not yet be evaluated. They added that, if successful, this approach to the treatment of cystic fibrosis would require periodic administration of gene therapy because trans-duction of an appropriate stem-cell population was not involved.

Recently published clinical protocols for gene therapy for cystic fibrosis include the following:

- Boucher and colleagues (1994), *Gene therapy for cystic fibrosis using E1-deleted adenovirus: a phase I trial in the nasal cavity*
- Crystal and colleagues (1995a), *Evaluation of repeat administration of a replication deficient, recom-binant adenovirus containing the normal cystic fibrosis transmembrane conductance regulator cDNA to the airways of individuals with cystic fibrosis*

- Crystal and colleagues (1995b), *A Phase I study, in cystic fibrosis patients, of the safety, toxicity, and biological efficacy of a single administration of a replication deficient, recombinant adenovirus carrying the cDNA of the normal cystic fibrosis transmembrane conductance regulator gene in the lung*
- Flotte (1996), *A Phase I study of an adeno-associated virus-CFTR gene vector in adult CF patients with mild lung disease.*

Publication trends

The annual numbers of references to gene therapy for cystic fibrosis identified through a search of the Medline database are shown *Figure 16.* The first reference appeared in 1987, after which there was a lull and no references were retrieved for 1988 and 1989. References to gene therapy for cystic fibrosis began to appear again in 1990, dipped slightly in 1991 and, since then, have been increasing stead-ily. Over 70 references were retrieved for 1995, although the number is still relatively small com-pared with the annual number references for lap-aroscopic cholecystectomy at its peak. The inform-ation in *Figure 16* appears to indicate that gene therapy for cystic fibrosis is still in the early phase of a continuing increase in publication trends.

Discussion

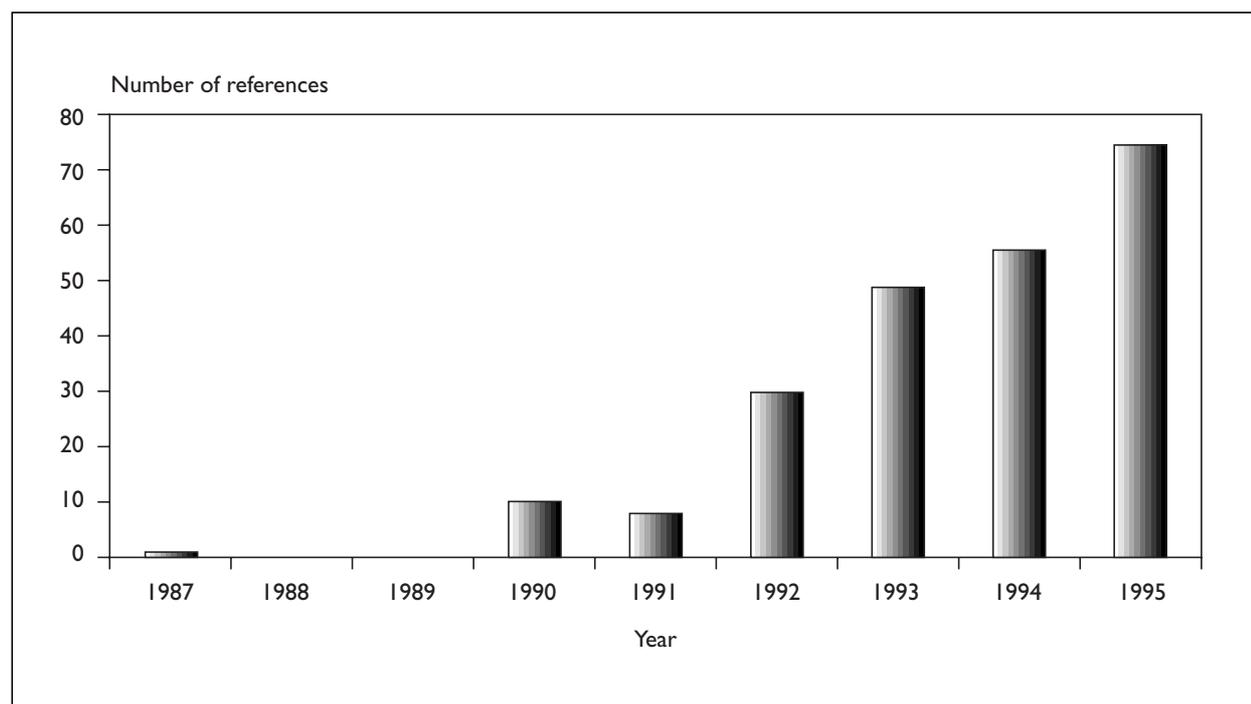In the USA, a panel of experts reporting on gene research to the National Institutes of

FIGURE 16 *Medline references for gene therapy for cystic fibrosis*

Health voiced concerns that the great potential of gene therapy might be undermined by efforts to rush it into clinical trials prematurely (Touchette, 1996). The panel pointed out that the field was in its infancy and that many problems had to be addressed before making major investments in human clinical trials. It also noted that most studies had neglected to include well-defined clinical end-points that would clearly indicate whether the therapy is having the desired effect. This has not discouraged academic and commercial enthusiasm for moving into clinical trials.

Main clinical approaches

The studies indicated that there are currently two main approaches to gene therapy for cystic fibrosis. Liposome-mediated gene transfer has been the vector system selected for initial clinical studies in the UK, while clinical trials in the USA are using adenovirus-mediated gene transfer. The problem of inflammation induced by the adenoviral vector would seem to be a special problem with cystic fibrosis patients, whose lungs are already damaged. Dorin (1996) commented that the alternative approach to gene delivery, using DNA liposome complexes, avoided most, if not all, of these immunological problems. It is still too early to predict whether either or both of these approaches will be successful.

Timing

Little or no information was provided in the reports about when the trials were undertaken or the reasons for their timing. It seems clear, however, that a major factor affecting timing is that the development of gene therapy is being controlled much like that of a therapeutic drug, and there exists a strict centralised regulatory process which clinical protocols have to pass through before approval to proceed is given. Thus, technical capability is determining the point at which trials are planned but the evolving regulatory processes are deciding the dates at which they are implemented. This slows the process of evaluation and allows time for reflection, which is not necessarily a bad thing where unconsidered and novel problems, both practical and ethical, are arising. It introduces yet another dimension into the question of the timing of assessments.

Impact of commercial interest

Pharmaceutical and biotechnology companies have invested large amounts of money in the expectation that gene-therapy treatments will eventually realise a commercial return; hence, there will be pressure from this sector to translate research findings into marketable products at the earliest opportunity. It will be informative to study how these companies attempt to bring pressure to bear on the regulators. In other cases, pharmaceutical companies have been skilful manipulators of popular opinion. This development will continue to offer an interesting case study for some time.

Chapter 9
Conclusions and recommendations

Summary

The primary objective of this systematic review was to try to identify the optimal time at which to initiate assessments of new and fast-changing health technologies. A series of literature searches were undertaken in an attempt to identify papers focusing on, first, the general principles involved in timing of HTAs and, second, reported assessments of six specific medical applications. Reported assessments of laparoscopic cholecystectomy, CVS, teleradiology, teledermatology, genetic screening for predisposition to breast cancer, and gene therapy for cystic fibrosis were analysed in order to try to identify factors that influenced the timing of those assessments. The six medical technologies offered a number of useful paired comparisons between those that were 'tele-' and 'chromosome/gene-' based, and those that were new and evolving, or relatively well-established. Since the literature relating to the six applications contained little information about the timing of the assessments, a number of interviews with key individuals were undertaken. This provided important insights but since this approach was not part of the original research plan, it was highly constrained by the time limit. The results indicated that future studies of this type should include more extensive use of carefully selected experts to supplement the inevitably limited information yielded by literature reviews in an area which has attracted little previous attention.

A bibliometric study of publication trends in the six applications was also undertaken, in an attempt to identify points in the development of a technology which could be used as indicators that assessment should be initiated. The results were inconclusive but further elaboration of the search strategy might yield useful, if crude, prospective indicators.

The general conclusions of the study were that assessment should be initiated early, using a variety of complementary assessment approaches. There were problems associated with all methods and, hence, a varied, empirical and iterative approach gave the most reliable results. Methods of assessment and reporting should be more standardised from the earliest stages, to improve the comparability of data. Resource issues should be incorporated into assessments from an early stage. Trials should be randomised from the outset. The precise point at which initiation of assessment should take place was not identified, either through the literature review or through the bibliometric study. However, the bibliometric study produced encouraging results. Further bibliometric research is recommended on a larger number of established technologies to detect whether there is a sufficiently consistent pattern to the publication trends of new and fast-changing health technologies to allow identification of a 'critical point' at which assessment should be recommended.

For all health technologies, more regular reporting of outcomes and side-effects should be encouraged during the period after assessment. In addition, as initial steps, guidelines should be established to standardise such reporting procedures and incentives developed to encourage adherence to such reporting procedures.

In areas where the technology is fast-changing, reassessment should take place from time to time. Again, the precise intervals were not identified and further research is needed to develop useful criteria for deciding when a technology has changed enough to warrant reassessment.

Conclusions

The general consensus that emerged from the review on timing was that:

- assessments should be initiated early
- assessments should use a variety of approaches to overcome some of the problems associated with each method
- assessments should be iterative.

Chalmers (1975) recommended randomisation of the first sick patient to receive a new drug or undergo a new procedure. He argued that it was probably unethical for a clinician to ask certain patients to forego their right to the standard accepted therapy and to be treated instead by a procedure not yet sufficiently developed to warrant comparison with the standard therapy. It was acknowledged that the result of such an

evaluation could prove to be less useful than expected, taking place as it was against a background of on-going modification of the procedure. The benefit lay in the fact that 50% of patients (those randomised to standard therapy) would not be subjected to the unknown risks of an unevaluated procedure. These issues are illustrated and discussed in the chapter on CVS.

In practice, the finite nature of resources for HTA means that it is not feasible to evaluate all new health technologies through RCTs. Black (1996) drew attention to circumstances when RCTs might be unnecessary, inappropriate, inadequate or even impossible. The Department of Health (1992) report also argued that observational studies may be sufficient when a technology has a large and clearly demonstrated impact.

While most assessment problems were exacerbated by a high rate of change in a technology, they were not qualitatively different. There were usually difficulties associated with trying to obtain all the data required for optimal assessments. Cost was a factor, but others were also important. For example, it emerged that ethical concerns could conflict with the requirements of rigorous assessment. RCTs might be terminated for 'ethical reasons' once sufficient clinical data had been collected to convince the investigator (of the usefulness or otherwise of the health technology), even though further trials might be needed to convince others or to gain enough economic data.

Another important point that arose in relation to economic impact evaluations was the difficulty of reliably predicting the pattern of demand when a new technology became available. This was particularly evident in the case of MAS, where substantial capital and training investment was already being made on the basis of predictions made a few years ago that most surgery would use this approach within the next few years. More recent evaluations of laparoscopic cholecystectomy gave less positive outcomes than earlier studies. This in turn could be expected to lead to greater conservatism in predictions of future demand. Telemedicine applications were also expected to change the pattern of demand, but the economic impact would depend on how service delivery was modified to deal with it. Thus economic evaluations must make assumptions that can turn out to be false and misleading.

The research questions posed in the Introduction are discussed below in the light of the findings reported earlier.

What general principles have been reported as guiding the timing of HTAs in the past?
Complexity of the pattern
Depending on their stage of development, health technologies can, in theory, be classified as future (not yet developed), emerging (prior to adoption), new (in the phase of adoption), accepted (in general use), and obsolete (superseded). However, in practice they do not always fit neatly into any one of these categories. Only highly-regulated technologies such as pharmaceuticals are constrained into this straightforward linear pattern with an orderly progression through the above phases. Over their life cycles, less-regulated health technologies develop and diffuse in a much more complex manner and it is not always possible to identify clear transition points between phases.

Factors affecting complexity
When development is locally regulated, or unregulated, innovators at different sites follow different protocols in parallel developments. This leads to variations of the application and non-comparable development data. It also leads to the discovery of other applications at some sites, followed by the pursuit of different development paths leading again to divergence of the applications and non-comparable data. This may restrict the possibility of accumulating sufficient rigorously controlled trial results on any one application to constitute adequate assessment.

Lack of regulation is not the only reason why health technologies diverge. Many now in use were originally developed for applications outside the health sector and have been incorporated through a process of technology transfer. Telemedicine applications fall into this category. The generic technology evolved in other industries in parallel with any use in healthcare, and innovations in these industries can be brought across into the healthcare sector at any time when an application is identified. This means that phases of the evolution of the technology are invisible or difficult to anticipate/predict for the healthcare sector and, consequently, not accessible for assessment.

Thus, health technologies evolve in a complex manner, which complicates the question of when to initiate assessment. There are also difficult value judgements to be made in fast-evolving technologies about when an application is new, and when changes are too slight to warrant the title 'new' that would trigger the need for an assessment.

Assessments of the six medical applications – factors influencing timing

General factors which influence why, rather than when, assessments are undertaken relate to the need to demonstrate that a new health technology is functionally capable, safe, works in practice, is cost-effective, value-added, and socially, legally and ethically acceptable.

Reviews of assessments of the six medical applications were undertaken in an attempt to clarify the factors that had influenced the timing of those assessments, and thus might prove relevant to the wider debate on timing. Published reports of assessments on the whole did not give specific reasons for their timing. Occasionally there were general statements explaining why it was consider-ed necessary to undertake such a trial, but more often the published report consisted of a straight-forward account of what had taken place. Never-theless, an analysis of reported assessments of the medical applications, and of existing reviews of the literature, when available, provided inform-ation on a number of factors that appear to have influenced the timing, directly or indirectly; these factors are summarised below.

Product champions and opinion leaders pioneer the introduction of new technologies into clinical practice and their observational reports may lead to a situation where more widespread diffusion of such technologies occurs, as was the case with laparoscopic cholecystectomy. Such novel proce-dures may diffuse rapidly before they can be ade-quately evaluated and this diffusion may place constraints on the method of evaluation that can be used. It is therefore important that assessments of new health technologies are undertaken before diffusion takes place.

The extent to which *regulatory control* is imposed on the introduction of new health technologies can also influence the timing of assessments. The presence of such controls might have helped to restrict the diffusion of laparoscopic cholecystec-tomy, making a large and widely generalisable RCT feasible.

The *source and level of funding* for studies may influence which aspects of a health technology are focused on and what type of trial is undertaken. Many telemedicine applications were funded by the commercial telecommunications organisations, who were perhaps more interested in proving that the technology worked and stimulating demand for their product than in demonstrating evidence-based benefits to patients. Short-term funding for

such projects may have led to the adoption of observational rather than experimental study designs, given the brief time-scale of such evaluations. The timing of assessments is also influenced by when technology providers make funding available.

Media coverage undoubtedly has an influence although it is difficult to predict exactly how this will operate and, hence, it must be treated with caution. One possible mode of influence is through generating favourable publicity about new, unevaluated health technologies, which can lead to immediate patient or physician demands for the new technique; such was the case for laparoscopic cholecystectomy, with its apparent benefits of a smaller scar, less postoperative pain, reduced hospital stay, and a more rapid recovery. The implications of this example for timing are that assessments need to be undertaken **before** such media coverage exerts popular pressure on purchasers to adopt the technology and dissuades patients from participating in RCTs in case they are randomised to the standard treatment. This, in turn, requires innovators to be more cautious in the claims that they make to the media. Random-isation had to be abandoned in a large US trial in which CVS was compared with amniocentesis because of difficulties with *patient recruitment* (Rhoads *et al,* 1989) resulting from inappropriate claims by doctors and subsequent public demand.

Another factor that affects the timing of assessments is *clinical uncertainty or equipoise,* the situation in which a clinician has no preference between the treatment options to be compared. As the period where a clinician has no preference for either a new or a standard treatment may be short, this has implications for the timing of assess-ments in the form of RCTs. Clinicians may be pre-pared to ask patients to participate in trials when they (the clinicians) are genuinely uncertain about which treatment is best. Once clinicians come to prefer either the standard or the alternative treat-ment, however, they may consider that, from an ethical viewpoint, they have a duty to provide only that treatment which they believe to be in the best interests of their patients. This argument was given as a reason for abandoning a proposed RCT of laparoscopic versus open cholecystectomy (Neuge-bauer *et al,* 1991). The counter argument, of course, is that randomisation is a hedging, risk-minimisation strategy when the true risks and benefits are not known.

The existence of the clinical *learning curve* also influences the timing of assessments of new

technologies. It has been suggested that assessments undertaken before clinicians have acquired sufficient skill in a new procedure will result in misleading findings on benefits and costs. Postponing an assessment until an appropriate point on the learning curve has been reached will largely dictate the timing of such an assessment. However, this appropriate time can usually only be recognised in retrospect and, when the stability of a new technology becomes obvious, clinicians may no longer be prepared to randomise.

Finally, the fact that the development of some health technologies is *technology-driven* or *commercially-driven* rather than needs-based, has impacted on timing to the extent that advances in the technology, in conjunction with a reduction in costs, have largely determined the timing of assessments, as, for example, in the case of teleradiology.

Are the methods in use adequate for new, fast-changing technologies?

Questions arose about whether the methods in use were adequate for any technologies. These came from a growing awareness that HTA had traditionally focused very narrowly on clinical outcomes. Now there were demands for a much wider range of criteria to be addressed in assessments, including social and ethical impact, effect on patterns of healthcare demand and other issues.

The reviews reported earlier make it clear that HTA can never be perfect under any circumstances. Best practice uses a number of methods of HTA, rigorously applied and reported, to achieve the most satisfactory outcomes for patients, but even in combination none of these methods is foolproof. Rare side-effects are often only detected after extensive use. New problems arise because of different ethical and cultural concerns in different patient groups. New patterns of demand created by the availability of new techniques can invalidate economic studies.

The problems with assessing the fast-changing technologies examined here were not qualitatively different, they were just likely to arise more frequently during the development phase of a technology. In the case of genetic technologies, regulation restricted them to research use until assessment gave satisfactory outcomes. Whether they continued to evolve after a technology had moved to routine use would be a function of the level of regulation of the application.

Cystic fibrosis gene therapy is being developed under tight central regulation comparable to that governing drugs. If approved, it would be used in an extremely constrained fashion. Any changes would have to be subjected to extensive testing. Genetic diagnostic testing for breast-cancer susceptibility, a non-invasive procedure, is subject to less tight regulation but there is tight control on diffusion. Telemedicine, on the other hand, is only beginning to be assessed on a limited basis, and there are no controls on adoption. This is a function of its low-risk profile in the eyes of users.

Thus the approaches to assessment were more a function of perceived risk than of rate of change. The methods in use were no less adequate for fast-changing technologies than for more stable technologies. What was lacking, for both stable and fast-evolving technologies, was a framework of standard guidelines and incentives to ensure that users would assess unregulated or lightly regulated health technologies in an approved and consistent fashion, and report the results. In addition, the guidelines now being issued for new surgical procedures need to be clarified to ensure that a decision about when a procedure has changed enough to be regarded as new is clearer and less subjective.

What gaps does a literature review reveal in the reporting of the principles that guide the timing of HTAs?

The literature provided no ready-made formula to guide the timing of HTAs but there was general agreement on some important principles. The assessment of new health technologies should be undertaken as early as is practicable, to provide timely information that aids decisions about whether the technology should be more widely adopted. As the technology is modified, or changes in competing technologies take place, re-assessment may be necessary. However, indicators to guide decisions about precisely when to initiate assessments were not offered, nor were suggestions made about how to deal with conflicts between the need for different types of assessment and ethical constraints.

A full review of the factors associated with timing would demand the use of a broader literature and a greater range of expertise than the scope of this study permitted. Future studies should take account of this in their methodology. Non-clinical issues, including the social and commercial issues touched on by some of the writers reviewed, have a bearing on decisions of timing. Hence, the literature on social science/social policy might also yield valuable insights.

How and when have assessments of the specific applications reviewed been carried out?

The timing of assessments has often reflected the growth of either clinical concern about side-effects (laparoscopic cholecystectomy, CVS), political concern about being seen to do something (cystic fibrosis gene therapy, genetic diagnosis) or growing popular demand/concern (genetic technologies). Sometimes clinical concern and popular demand have come together, as with laparoscopic cholecystectomy where the public were demanding what they believed to be a better technique while clinicians were concerned that adverse side-effects were raising questions about the safety of the technique. In some of these cases, such as laparoscopic cholecystectomy, this meant that assessment was not initiated until it proved to be too late to carry out a large-scale RCT comparing laparoscopic with open cholecystectomy.

Laparoscopic cholecystectomy also illustrates the difficulty of achieving properly controlled trials when public interest has been aroused and people have formed a consensus view about a procedure. Thus media coverage of a health technology, and/or physician influence, may engender patient reluctance to participate in an RCT. The need to acquire new skills also complicates the issue here. There is a real possibility that, by the time clinicians feel sufficiently skilled in a new procedure, conducting an RCT may no longer be a practical option and, in any case, practitioner preference may have been established with a loss of equipoise. Meanwhile, during the learning period, patients will have continued to be subjected to an unevaluated procedure, the performance of which is thought to be suboptimal.

In the case of cystic fibrosis gene therapy and CVS, an early decision to assess allowed effective control and thorough assessment approaches to be used. However, the highly regulated cystic fibrosis trials now under way are costly and slow. This highly precautionary approach is acceptable for invasive or otherwise risky applications but it may be difficult to justify imposing such regulation on innovations which are perceived by neither clinicians nor the public to have adverse potential. Where an application is neither invasive nor seen to have alarming implications for non-health-related aspects of the patient's life, it would be hard to impose expensive regulatory control. Telemedicine applications, for instance, would currently be unlikely to arouse sufficient concern. However, this laissez-faire approach may prove to be misguided if, for example, misdiagnosis using telemedicine leads to harm and successful litigation. In contrast to telemedicine, genetic testing, with its implications for life insurance and other important matters is widely perceived as requiring some regulation and control.

Thus public and clinical perception of the risks associated with an application are very important in determining whether assessment is undertaken early enough to be effective. This argues for greater education and debate about these issues.

Have opportunities for practical and desirable assessments been lost?

The continuing publication of data about laparoscopic cholecystectomy that contradicts some of the earlier findings confirms that it would have been desirable to have had a thorough and rigorous assessment of this application at an early stage. By waiting until equipoise had been lost before initiating assessment, the opportunity to obtain clear and unambiguous data was missed. This has complicated assessment approaches and has probably increased the costs of assessment and rendered the results less convincing.

Are there lessons to be learnt from the applications that have now reached a relatively stable state, which could be applied to the fast-evolving applications?

When assessment was initiated early and systematically, as with CVS, the initial study and subsequent work have provided a more useful and comprehensible body of information to guide decisions. When assessment was initiated late, after some diffusion had taken place, as with laparoscopic cholecystectomy, the studies were less satisfactory and rigorous, and the interpretation of data was not easy. This confirms the findings of the timing review, that assessment should be initiated early, before diffusion has taken place and before practitioners have developed preferences. However, while this is difficult to disagree with in theory, in practice the requirement for clinical uncertainty is extremely hard to ensure throughout trials. Inevitably, with use the practitioner tends to form personal views about the effectiveness of a procedure, with loss of the equipoise required for trial participation. Again, this argues for greater education of clinicians about HTA and the dangers of unsupported clinical impressions.

Another point which became evident with both of the 'stable' technologies was that although the rate of publication of evaluations had fallen off, studies were still being published which either conflicted with earlier studies (laparoscopic cholecystectomy)

or else added important data (CVS). They were not, in fact, entirely stable, although the rate of development had fallen to a low level. It is probable that, as in other industries, health technologies continue to evolve, although more slowly, until they are clearly superseded.

The reasons for near-stability were also different. Laparoscopic cholecystectomy had been widely accepted as a beneficial technique with an accepted protocol for safe use. In contrast, the use of CVS contracted after reports of rare serious side-effects. Its use was restricted to higher-risk pregnancies with the procedure performed only relatively late in the first trimester of pregnancy. Although it still has a place in clinical practice, it cannot be considered completely safe.

The main lesson to be learnt, then, is that already enunciated – assessment should be initiated early, should be an iterative process, and reporting of outcomes should continue throughout the life of a technology. In addition, a technology should be restricted to research use when a satisfactory, safe and reliable protocol for its use has not yet been established.

To what degree are assessments comparable at different points in the evolution of a fast-evolving technology?

The answer to this question depends on the nature of the assessment and the extent to which practitioners attempt to make their studies comparable with previous ones. The reviews demonstrated that studies have often failed to control the parameters of their trials in ways that permit comparability. This has often been for reasons beyond the control of the researchers, such as differences in the contexts of the studies which have impacted on outcomes. In other cases, greater comparability could have been achieved but was not attempted. The type of evaluation that is appropriate may change over time, depending on the parameters being assessed. Concerns about rare effects of CVS led to case–control studies, mostly after the trials had been concluded.

Can inferences be drawn about trends in development/diffusion of technologies, from collective characteristics of reported assessments, that can be used to decide on the timing of assessments?

The bibliometric studies covered too few applications to establish whether this approach might yield useful indicator points to trigger HTAs. In addition, some of the applications, such as teledermatology have such a limited

literature that the absolute numbers of publications were too small to use in this way. These are mainly applications used in primary care, which are not usually the subject of many research papers. Thus bibliometric trends may be particularly unhelpful as a source of indicators for primary care innovations. However, the publication trend curves for laparoscopic cholecystectomy and CVS gave more promising inflection points at times that approximated to dates at which assessment should have been initiated. This is suggestive, although the results are too limited to make generalisations. It is recommended that a study of at least ten further applications which have reached relative stability, but which are fairly recent, should be carried out. One limitation of this approach is that the reports of research are published at varying times after initiation.

Is there information that can be used to derive useful general guidelines and protocols for assessment? How much of a problem is posed by issues specific to the technology and application?

One conclusion from the study is that it would be helpful to categorise technologies into, for instance, diagnostic versus therapeutic, and invasive versus non-invasive. It would be more feasible to provide a set of protocols for each category than to try to derive a very general code applicable to all. The form of such protocols could be a flow diagram, containing a series of questions and directions, to assist purchasers in deciding whether they should undertake assessment or should refer to appropriate national bodies for guidance on the form assessment should take and the reporting rules. Such protocols could only be created after agreement on standard assessment and reporting procedures was reached.

Standardisation would be doubly useful because it would facilitate meta-analysis of studies. It emerges from the literature reviewed that there has been little attempt, in many cases, to make studies comparable with previous ones. If general guidelines were prepared, it would be important to set standards for assessments and reporting and to provide appropriate and effective incentives (and resources) to comply.

Does the information gleaned from the reviews yield insights as to when experimental services should move to routine status?

It is clear that services should not move to routine status until they have been well-assessed, training and equipment needs identified, and other aspects of evaluation carried out. The case of laparoscopic

cholecystectomy illustrates the importance of having this infrastructure in place before diffusion. The early evaluation of CVS led to its diffusion being restrained. Although early trials did demonstrate an increased risk of miscarriage, it was observational studies, using data collected after it had begun to enter routine practice, that suggested the rare adverse side-effects which had the greatest impact on practice. This underlines the fact that continuing surveillance may be as important as early assessment, and that decisions to move to routine status, based on HTAs, do not guarantee the safety of a technique. However, the quality and systematic nature of the original studies assisted the interpretation of the findings of all studies. Thus quality and timing are related, as suggested in chapter 1.

Modifications to the original protocol

The initial research methodology was modified by the addition of interviews with key individuals associated with applications development. From discussions within the research team and from the early stages of the review it became evident that important factors influencing timing and choice of method were known to individuals who were actively involved in the research, but were not recorded in the literature. This was a reminder that scientific articles have a convention of their own about the type of information reported and the format, which does not include information of the nature sought in this study. In future studies it is important to take these conventions into account when considering the role of a literature review in a research project. In the present case, the study would have been enhanced by an initial round of interviews with a wider range of key individuals associated with a larger number of applications, and by reducing the scope of the systematic reviews.

Recommendations for future research

As noted above, a further series of bibliometric studies is recommended in order to explore the possibility of using inflections in trend curves as indicators of when to initiate assessment. A wider range of literature could also usefully be scanned. For example, both in order to identify key indicator points and also to gain increased insight into the role of the media, women's media could be included in the range of literature reviewed and scanned. In this study, the role of consumer pressure was highlighted in the cases of CVS and breast-cancer diagnosis. It appears probable that important information could be gained about a number of other conditions in this way.

Another useful programme of research would be to create a typology of health technologies which would subdivide them into categories that could be subjected to standard assessment protocols, to identify standards for assessment and reporting for each category, and to consider the question of the influence of incentives and barriers in adhering to rigorous standards in HTAs.

The commentators all experienced difficulty in trying to achieve greater specificity about when, how, and how frequently to evaluate new health technologies. One reason for this may be that, with the exception of the bodies addressing the problems of surgical techniques, there has been a tendency to aggregate the whole range of new technologies (other than drugs) which are used to promote health, prevent and treat disease, and improve rehabilitation and long-term care. The literature gives little assistance with formulating a global theory. Rather it illustrates the very different characteristics of different healthcare technologies and the problems associated with a single methodological approach. Given their diversity it is unlikely that one approach will be suited to all, although there are clearly some important common principles. However, these relate more to the demands for relative safety and efficacy, than to a stereotyped protocol for the timing or nature of assessment. These demands might be met more satisfactorily by abandoning the search for a single theory in favour of adopting several distinct approaches.

A first step towards addressing this issue would be to reduce the complexity of any one case somewhat by categorising health technologies on the basis, for example, of their invasiveness and of possible characteristics such as their apparent cost advantages/disadvantages, their potential for major improvements to public health, or their ethical impact. A strongly precautionary approach would clearly be appropriate for the more invasive, ethically problematic and costly technologies. A health technology might be assigned a score on each of several parameters, and the overall total used to decide on priorities for evaluation. A series of different tests could be applied depending on the initial score, to decide on the next step. Provided the guidelines were sufficiently clear, this would allow decentralised care purchasers and ethics committees to decide on whether they were competent to evaluate a new technology, or whether they should refer to central authority for further guidance.

The question of which methods of evaluation to apply and when to apply them would be easier

to answer when the category of health technology was more precisely defined and its risk characteristics assessed. However, the general principle advocated in most of the literature reviewed here, of using a number of methods of evaluation on an iterative basis is clearly important, and its utility would be enhanced if a more unified standard of reporting in observational and experimental studies were to be observed. For example, the recent CONSORT initiative, reported by Begg (1996), is an attempt to improve the overall standard of reporting of RCTs. Hence a set of strong recommendations about how to carry out and record each type of evaluation would enhance the value of each. In addition, there is clearly a need to improve the effectiveness of current systems of reporting side-effects of treatments after they have become routine and the level of interest in the technologies has fallen.

Another implication for policy makers is that significant new health technologies need to be identified as they emerge, possibly through some form of horizon-scanning mechanism, in order that they can be prioritised and steps taken to assess them in an appropriate manner as early as is practicable. Otherwise, the danger remains that a combination of technology 'push', pioneer enthusiasm, and media publicity may propel the diffusion of new technologies into routine practice before they have been adequately evaluated. The implications for healthcare managers and clinicians are that, where the introduction of new technologies is not covered by central regulation, national initiatives such as SERNIP should be supported. In areas of medicine where no such national initiatives exist, the establishment of local guidelines would help to ensure that new health technologies were considered systematically before a decision was taken to adopt them. Healthcare managers and clinicians also need to be aware of the results of assessments which may already have been carried out on new health technologies whose introduction they are considering.

Acknowledgements

◆

References

A triumph of hype over experience? Gene therapy has promised much. It has yet to deliver the goods. *Economist* 1995;**337**(7945):111–12.

All our yesterdays. *Health Serv J* 1996;**106**:25.

Detectives on the case – early detection of cancer means a higher survival rate for patients. *Financial Times* 1995 January 24;13.

'Distance health care' is latest medicine. *Electronic Eng Times* 1996;**29**:55.

Genetic testing for cancer risk: research projects being funded. *Oncology* 1994;**8**:16–18.

Is my baby all right? Fetal diagnostic test chorionic villus sampling implicated in birth defects. *Newsweek* 1992 June 22;62–3.

Safer tests expected – a new method of prenatal diagnosis. *Financial Times* 1996 January 12;10.

Surgical innovation under scrutiny. *Lancet* 1993;**342**:187–8.

Telemedicine: big sister is watching you. *Economist* 1997;**342**(7999):51.

Telemedicine: fad or future? *Lancet* 1995;**345**:73–4.

Abrams H, Hessel S, 1987. Health technology assessment: problems and challenges. *Am J Roentgenol*;**149**:1127–32.

Abramsky L, Rodeck CH, 1991. Women's choices for fetal chromosome analysis. *Prenat Diagn*;**11**:23–8.

ACOST, 1993. A report on medical research and health. Advisory Council on Science and Technology. London: HMSO.

Adams ME, 1992. Economic analysis in randomized control trials. *Med Care*;**30**:231–8.

Adang EM, Dirksen CD, Engel GL, Baeten CG, 1995. Medical technology assessment: economic evaluation of new technologies. *Br J Hosp Med*;**53**:563–6.

Agnifili A, Ibi I, Guadagni S, *et al.,* 1993. Perioperative pain and stress: a comparison between video laparoscopic cholecystectomy and "open" cholecystectomy. *G Chir*;**14**:344–8.

Alder, H,1988. Organization key to technology diffusion. *Hospitals*;**62**(13):82FB.

Alfirevic Z, Gosden C, Neilson JP, 1995a. Chorion villus sampling compared with amniocentesis for prenatal diagnosis. Cochrane Database of Systematic Reviews.

Alfirevic Z, Gosden C, Neilson JP, 1995b. Early amniocentesis vs chorion villus sampling. Cochrane Database of Systematic Reviews.

Andersson F, 1995. Why is the pharmaceutical industry investing increasing amounts in health economic evaluations? *Int J Technol Assess Health Care*;**11**:750–61.

Andrade JD, 1991. 'What, me worry?' Bioengineering and the costs of health care. *Ann Biomed Eng*;**19**:550–1.

Antczak-Bouckoms A, Burdick E, Klawansky S, Mosteller F, 1991. Using medical registries and data sets for technology assessment. *Int J Technol Assess Health Care*;**7**:123–8.

Bailey RW, Zucker KA, Flowers JL, *et al.,* 1991. Laparoscopic cholecystectomy. Experience with 375 consecutive patients. *Ann Surg*;**214**:531–41.

Ball K, 1995. Can we afford advanced and sophisticated technology in tomorrow's health care arena? *Minim Invasive Surg Nurs*;**9**:53.

Banta HD, 1986. Dutch committee assesses the future of health technology. *Dimens Health Serv*;**63**:17–20.

Banta HD, 1990. The regulation of medical devices. *Prev Med*;**19**:693–9.

Banta HD, 1993. Minimally invasive therapy in five European countries: diffusion, effectiveness and cost-effectiveness. *Health Policy*;**23**:1–178.

Banta HD, Andreasen PB, 1990. The political dimension in health care technology assessment programs. *Int J Technol Assess Health Care*;**6**:115–23.

Banta HD, Gelijns AC, 1994. The future and health care technology: implications of a system for early identification. *World Health Stat*;**47**:140–8.

Banta HD, Luce BR, 1993. Health care technology and its assessment: an international perspective. Oxford: Oxford University Press.

Banta HD, Thacker SB, 1990. The case for reassessment of health care technology. Once is not enough. *JAMA*;**264**:235–40.

Banta HD, van Beekum WT, 1990. The regulation of medical devices and quality of medical care. *Qual Assur Health Care*;**2**:127–36.

Banta HD, Schersten T, Jonsson E, 1993. Implications of minimally invasive therapy. *Health Policy*;**23**:167–78.

Barkun JS, Barkun AN, Sampalis JS, *et al.,* 1992. Randomised controlled trial of laparoscopic versus mini cholecystectomy. *Lancet*;**340**:1116–19.

Barkun JS, Caro JJ, Barkun AN, Trindade E, 1995. Cost-effectiveness of laparoscopic and mini-cholecystectomy in a prospective randomized trial. *Surg Endosc*;**9**:1221–4.

Barondess JA, 1994. Rational use of new technologies. *Health Aff (Millwood)*;**13**:275–6.

Bashshur RL, 1995. Telemedicine effects: cost, quality, and access. *J Med Syst*;**19**:81–91.

Bashshur R, Lovett J, 1977. Assessment of telemedicine: results of the initial experience. *Aviat Space Environ Med*;**48**:65–70.

Bashshur RL, Armstrong PA, Youssef ZI, 1975. Telemedicine: explorations in the use of telecommunications in health care. Springfield, VA: Charles C Thomas.

Bass EB, Pitt HA, Lillemoe KD, 1993. Cost-effectiveness of laparoscopic cholecystectomy versus open cholecystectomy. *Am J Surg*;**165**:466–71.

Battista RN, Hodge MJ, 1995. The development of health care technology assessment. An international perspective. *Int J Technol Assess Health Care*;**11**:287–300.

Battista RN, Feeny DH, Hodge MJ, 1995. Evaluation of the Canadian Coordinating Office for Health Technology Assessment. *Int J Technol Assess Health Care*;**11**:102–16.

Baxter JN, O'Dwyer PJ, 1992. Laparoscopic or minilaparotomy cholecystectomy? *BMJ*;**304**:559–60.

Beardsley T, 1996. Vital data. *Sci Am*;**274**:76–81.

Begg C, Cho M, Eastwood S *et al.,* 1996. Improving the quality of reporting of randomized controlled trials: the CONSORT statement. *JAMA*;**276**:637–9.

Berci G, Sackier JM, 1991. The Los Angeles experience with laparoscopic cholecystectomy. *Am J Surg*;**161**:382–4.

Berg AO, 1992. Assessing a technology: what constitutes enough? *J Am Board Fam Pract*;**5**:439–40.

Bergevin Y, Tugwell P, 1995. Introduction: needs-based technology assessment. Who can afford not to use it? *Int J Technol Assess Health Care*;**11**:647–9.

Berggren U, Gordh T, Grama D, Haglund U, Rastad J, Arvidsson D, 1994. Laparoscopic versus open cholecystectomy: Hospitalization, sick leave, analgesia and trauma responses. *Br J Surg*;**81**:1362–5.

Bergman R, 1993. Letting telemedicine do the walking. *Hosp Health Netw*;**67**:46–8.

Black N, 1996. Why we need observational studies to evaluate the effectiveness of health care. *BMJ*;**312**:1215–18.

Blais R, 1991. Using administrative data bases for technology assessment in health care. *Int J Technol Assess Health Care*;**7**:203–8.

Blakemore KJ, 1988. Prenatal diagnosis by chorionic villus sampling. *Obstet Gynecol Clin North Am*;**15**:179–213.

Blanpain JE, 1984. Health technology assessment in Belgium, France, Germany and Japan. *Acta Hospitalia*;(3):37–44.

Bloor K, Maynard A, 1994. Through the keyhole. *Health Serv J*;(17 November):24–6.

Bond S, 1994. What are nursing's priorities? *Nurs Standard*;**8**(38):20–1.

Border P, 1995. Minimal access ('keyhole') surgery and its implications. London: Parliamentary Office of Science and Technology.

Bouchard S, Barkun AN, Barkun JS, Joseph L, 1996. Technology assessment in laparoscopic general surgery and gastrointestinal endoscopy: science or convenience? *Gastroenterology*;**110**:915–25.

Boucher RC, Knowles MR, Johnson LG, *et al.,* 1994. Gene therapy for cystic fibrosis using E1-deleted adenovirus: a phase I trial in the nasal cavity. *Hum Gene Ther*;**5**:615–39.

Bower DJ, 1994. Innovation in health care: developments in NHS trusts. *J Manage Med*;**8**:54–61.

Brambati B, 1995. Chorionic villus sampling. *Curr Opin Obstet Gynecol*;**7**:109–16.

Brambati B, Terzian E, Tognoni G, 1991. Randomized clinical trial of transabdominal versus transcervical chorionic villus sampling methods. *Prenat Diagn*;**11**:285–93.

Briggs A, Sculpher M, Buxton M, 1994. Uncertainty in the economic evaluation of health care technologies: the role of sensitivity analysis. *Health Econ*;**3**:95–104.

Brown KS, 1996. Investigators grappling with looming disease gene concerns. *Scientist*;**10**:1–9.

Brown N, Lumley J, Tickle C, Keene J, 1996. Congenital limb reduction defects. Clues from developmental biology, teratology and epidemiology. London: HMSO.

Bryant H, 1996. Genetic screening for breast cancer in Ashkenazi women. *Lancet*;**347**:1638.

Bryce RL, Bradley MT, McCormick SM, 1989. To what extent would women prefer chorionic villus sampling to amniocentesis for prenatal diagnosis? *Paediatr Perinat Epidemiol*;**3**:137–45.

Bunker JP, 1981. Health care technology assessment essential to effective medical care. *Med Instrum*;**15**:289–90.

Bunker JP, Hinkley D, McDermott WV, 1978. Surgical innovation and its evaluation. *Science*;**200**:937–41.

Burke G, 1994. High-tech and health reform: the folly of doing it. *Hosp Health Netw*;**68**:120.

Burton BK, Schulz CJ, Burd LI, 1992. Limb anomalies associated with chorionic villus sampling. *Obstet Gynecol*;**79**:726–30.

Buto KA, 1994. How can Medicare keep pace with cutting-edge technology? *Health Aff (Millwood)*;**13**(Summer):137–40.

Byrne D, Marks K, Azar G, Nicolaides KH, 1991. Randomized study of early amniocentesis versus chorionic villus sampling: a technical and cytogenetic comparison of 650 patients. *Ultrasound Obstet Gynecol*;**1**:235–40.

Byrne J, Timon D, Armstrong C, Horgan PG, Quill DS, 1994. Comparison of analgesic requirements and pulmonary function in 'open' versus laparoscopic cholecystectomy. *Minim Invasive Ther*;**3**:3–6.

Canadian Collaborative CVS–Amniocentesis Clinical Trial Group, 1989. Multicentre randomised clinical trial of chorion villus sampling and amniocentesis. *Lancet*;**i**:1–6.

Caplen NJ, Gao X, Hayes P, *et al.,* 1994. Gene therapy for cystic fibrosis in humans by liposome-mediated DNA transfer: the production of resources and the regulatory process. *Gene Ther*;**1**:139–47.

Caplen NJ, Alton EW, Middleton PG, *et al.,* 1995. Liposome-mediated CFTR gene transfer to the nasal epithelium of patients with cystic fibrosis. *Nature Med*;**1**:39–46.

Challah S, Mays NB, 1986. Randomised controlled trial in the evaluation of new technology: a case study. *BMJ*;**292**:877–9.

Chalmers TC, 1975. Randomization of the first patient. *Med Clin North Am*;**59**:1035–8.

Chalmers I, 1997. What is the prior probability of a proposed new treatment being superior to established treatments? *BMJ*;**314**:74–5.

Chen PS Jr, 1993. Biotechnology transfer at the National Institutes of Health. *Qual Assur*;**2**:345–52.

Chitty LS, Bobrow M, 1994. Techniques in current use in prenatal diagnosis. *Int J Technol Assess Health Care*;**10**:604–27.

Christensen K, 1996. Novel genetics research and technologies present major challenges for public policy. *Genet Eng News*;**16**:4–35.

Clarysse B, Debackere K, Rappa MA, 1996. Modelling the persistence of organizations in an emerging field: the case of hepatitis C. *Res Policy*;**25**:671–88.

Clothier C, 1992. Report of the Committee on the Ethics of Gene Therapy. London: HMSO: Cm 1788.

Coelho JC, de Araujo RP, Marchesini JB, Coelho IC, de Araujo LR, 1993. Pulmonary function after cholecystectomy performed through Kocher's incision, a mini-incision, and laparoscopy. *World J Surg*;**17**:544–6.

Coles SFS, 1995. Telemedicine: the rise of digital healthcare. London: FT Pharmaceuticals & Healthcare Publishing.

Comroe JH, Dripps RD, 1981. Scientific basis for the support of biomedical science. In: Roberts EB, Levy RI, Finkelstein SN, Moskowitz J, Sondik EJ, editors. Biomedical innovation. Cambridge, Mass: MIT Press:101–22.

Coombs R, Narandren P, Richards A, 1996. A literature-based innovation output indicator. *Res Policy*;**25**:403–13.

Cram N, Wheeler J, Lessard CS, 1995. Ethical issues of life-sustaining technology. *IEEE Technol Soc Mag*;**14**:21–8.

Crepea AT, 1995. A systems engineering approach to technology assessment. *J Clin Eng*;**20**:297–303.

Crump WJ, Pfeil T, 1995. A telemedicine primer. An introduction to the technology and an overview of the literature. *Arch Fam Med*;**4**:796–803.

Crystal RG, McElvaney NG, Rosenfeld MA, *et al.,* 1994. Administration of an adenovirus containing the human CFTR cDNA to the respiratory tract of individuals with cystic fibrosis. *Nat Genet*;**8**:42–51.

Crystal RG, Mastrangeli A, Sanders A, *et al.,* 1995a. Evaluation of repeat administration of a replication deficient, recombinant adenovirus containing the normal cystic fibrosis transmembrane conductance regulator cDNA to the airways of individuals with cystic fibrosis. *Hum Gene Ther*;**6**:667–703.

Crystal RG, Jaffe A, Brody S, *et al.,* 1995b. A phase 1 study, in cystic fibrosis patients, of the safety, toxicity, and bio-logical efficacy of a single administration of a replication deficient, recombinant adenovirus carrying the cDNA of the normal cystic fibrosis transmembrane conductance regulator gene in the lung. *Hum Gene Ther*;**6**:643–66.

Culyer AJ, 1988. Technology assessment in Europe: its present and future roles. In: Rutten FFH, Reiser ST, editors. Economics of medical technology. Berlin: Springer-Verlag.

Culyer A, 1994. Supporting research and development in the NHS. London: HMSO.

Cuschieri A, 1989. The laparoscopic revolution – walk carefully before we run. *J R Coll Surg Edinb*;**34**:295.

Cuschieri A, 1994. Minimal access surgery: implications for the NHS. Report from a Working Group chaired by Professor Alfred Cuschieri. Edinburgh: HMSO.

Cuschieri A, 1995. Whither minimal access surgery: tribulations and expectations. *Am J Surg*;**169**:9–19.

Cuschieri A, Dubois F, Mouiel J, *et al.,* 1991. The European experience with laparoscopic cholecystectomy. *Am J Surg*;**161**:385–7.

David Y, 1989. Technology related decision-making issues in hospitals. *Proc Annual Conf Eng Med Biol*; **11**(Pt 5):1553–4.

David Y, 1993. Technology evaluation in a US hospital: the role of clinical engineering. *Med Biol Eng Comput*;**31**:28–32.

David Y, Cheng S, Natale D, Whitson H, Keller J, Larkin B, 1993. New approaches to technology assessment: opportunities and trends. *Biomed Instrum Technol*;**27**:371–6.

Davidoff AJ, Powe NR, 1996. Role of perspective in defining economic measures for the evaluation of medical technology. *Int J Technol Assess Health Care*;**12**:9–21.

Davies L, Coyle D, Drummond M, 1994. Current status of economic appraisal of health technology in the European Community: report of the network. (The EC Network on the Methodology of Economic Appraisal of Health Technology.) *Soc Sci Med*;**38**:1601–7.

Davis C, Arregani M, Nagan RF, *et al.,* 1992. Laparoscopic cholecystectomy: the St Vincent experience. *Surg Laparosc Endosc,***2**:64–9.

Deber RB, 1992. Translating technology assessment into policy. Conceptual issues and tough choices. *Int J Technol Assess Health Care,***8**:131–7.

de Charro FT, 1990. The economics of technological change in health. The conditions governing the production and diffusion of innovations in health care. *Biotherapy,***2**:271–4.

DeMeyer A, Saab W, Tabet J, 1994. The development of scientific networks and multinational research projects: who plays the key role? In: Managing human resources in R&D. Proceedings of the 1994 R&D Management Conference; 1994 July 4–6; Manchester. Manchester Business School.

Department of Health, 1992. Assessing the effects of health technologies: principles, practice, proposals. London: Department of Health.

Department of Health, 1995. Proposed safety and efficacy register of new interventional procedures of the Medical Royal Colleges. London: Department of Health.

Diamond GA, Denton TA, 1993. Alternative perspectives on the biased foundations of medical technology assessment. *Ann Intern Med,***118**:455–64.

Dickersin K, Herxheimer A, 1996. Introduction: the quality of the medical evidence: is it good enough? *Int J Technol Assess Health Care,***12**:187–9.

DiMasi JA, 1991. Cost of innovation in the pharmaceutical industry. *J Health Econ,***10**:107–42.

Dolan A, Zingg W, 1993. Health care technology: how can we tell if we can afford it? A Canadian viewpoint. *J Long Term Effects Med Implants,***3**:277–82.

Donaldson MS, Sox HC, 1992. Setting priorities for health technology assessment: a model process. Washington DC: National Academy Press.

Dorin J, 1996. Somatic gene therapy: optimism tempered by reality. *BMJ,***312**:323–4.

Downs SH, Black NA, Devlin HB, Royston CMS, Russell RCG, 1996. Systematic review of the effectiveness and safety of laparoscopic cholecystectomy. *Ann R Coll Surg Eng,***78**:235–323.

Drummond MF, 1990. Allocating resources. *Int J Technol Assess Health Care,***6**:77–92.

Drummond M, 1992. Test drive. *Health Serv J,***102**:26–7.

Drummond M, 1994. Evaluation of health technology: economic issues for health policy and policy issues for economic appraisal. *Soc Sci Med,***38**:1593–1600.

Drummond, MF, Bloom BS, Carrin G, *et al.,* 1992. Issues in the cross-national assessment of health technology. *Int J Technol Assess Health Care,***8**:671–82.

Dubois F, Berthelot G, Levard H, 1991. Laparoscopic cholecystectomy: historic perspective and personal experience. *Surg Laparosc Endosc,***1**:52–7.

Durand-Zaleski I, Jolly D, 1990. Technology assessment in health care – decision makers and health care providers: what they need to know. *Health Policy,***15**:37–44.

Eddy DM, 1990. Should we change the rules for evaluating medical technologies? In: Gelijns AC, editor. Modern methods of clinical investigation. Washington DC: National Academy Press.

Edwards KL, Gordon TJ, 1984. Characterisations of innovations introduced on the US market since 1992. Glastonbury, Conn: Futures Group.

Eeles R, 1996. Testing for the breast cancer predisposition gene, BRCA1. *BMJ,***313**:572.

Elliott D, Hollins B, 1995. Product evaluation: theoretical and practical considerations. *Aust Crit Care,***8**:14–19.

European Working Group on Human Gene Transfer and Therapy Central Office, 1996. Human gene therapy clinical trials in Europe. *Hum Gene Ther,***7**:1258–9.

Evans DG, 1995. Genetic testing for cancer predisposition: need and demand. *J Med Genet,***32**:161.

Evans JA, Hamerton JL, 1996. Limb defects and chorionic villus sampling. *Lancet,***347**:484–5.

Eyre H, Sondik E, Smith RA, Kessler L, 1995. Joint meeting on the feasibility of a study of screening premenopausal women (40–49 years) for breast cancer: April 20–21, 1994. *Cancer,***75**:1391–403.

Fahy MJ, Lippman A, 1988. Prenatal diagnosis and the Canadian Collaborative Randomized Trial of chorionic villi sampling: the physician's view. *Am J Med Genet,***31**:953–61.

Feeny DH, Torrance GW, 1989. Incorporating utility-based quality-of-life assessment measures in clinical trials. Two examples. *Med Care,***27**(3 suppl):190–204.

Feeny D, Guyatt G, Tugwell P, 1986. Health care technology: effectiveness, efficiency and public policy. Canada: Institute for Research on Public Policy.

Ferguson JH, Dubinsky M, Kirsch PJ, 1993. Court-ordered reimbursement for unproven medical technology: Circumventing technology assessment. *JAMA,***269**:2116–21.

Ferzli G, Kloss DA, 1991. Laparoscopic cholecystectomy: 111 consecutive cases. *Am J Gastroenterol,***86**:1176–8.

Fineberg HV, 1985. Technology assessment. Motivation, capability, and future directions. *Med Care,***23**:663–71.

Firshein J, 1986. HCFA urged to speed approval of technologies. *Hospitals,***60**:67.

Firth HV, Boyd PA, Chamberlain P, MacKenzie IZ, Lindenbaum RH, Huson SM, 1991. Severe limb abnormalities after chorion villus sampling at 56–66 days' gestation. *Lancet,***337**:762–3.

Flotte TE, 1996. Clinical protocol. A phase I study of an adeno-associated virus-CFTR gene vector in adult CF patients with mild lung disease. *Hum Gene Ther,***7**:1145–59.

Franklin C, 1993. Basic concepts and fundamental issues in technology assessment. *Intensive Care Med*;**19**:117–21.

Franklin C, 1994. Effectiveness and efficacy – which is which? (1). *Crit Care Med*;**22**:1335–6.

Froster UG, Jackson L, 1996. Limb defects and chorionic villus sampling: results from an international registry, 1992–94. *Lancet*;**347**:489–94.

Fuchs M, 1979. Provider attitudes toward STARPAHC: a telemedicine project on the Papago Reservation. *Med Care*;**17**:59–68.

Fuchs VR, Garber AM, 1990. The new technology assessment. *N Engl J Med*;**323**:673–7.

Fullarton GM, 1992. Evaluation of the cost of laparoscopic and open cholecystectomy. *Gut*;**33**:17S.

Gafni A, Birch S, 1993. Guidelines for the adoption of new technologies: a prescription for uncontrolled growth in expenditures and how to avoid the problem. *Can Med Assoc J*;**148**:913–17.

Garber AM, 1994. Can technology assessment control health spending? *Health Aff (Millwood)*;**13**:115–26.

Garcia-Caballero M, Vara-Thorbeck C, 1993. The evolution of postoperative ileus after laparoscopic cholecystectomy. A comparative study with conventional cholecystectomy and sympathetic blockade treatment. *Surg Endosc*;**7**:416–19.

Garud RA, Rappa MA, 1992. On the persistence of researchers in technological development. In: Best papers. Proceedings of the 52nd Annual Meeting of the American Academy of Management. Wall & Jauch: 369–73.

Gayler BW, Gitlin JN, Rappaport W, Skinner FL, Cerva J, 1981. Teleradiology: an evaluation of a microcomputer-based system. *Radiology*;**140**:355–60.

Gelijns AC, 1990. Modern methods of clinical investigation. Washington DC: National Academy Press.

Gelijns AC, Fendrick AM, 1993. Dynamics of innovation in minimally invasive therapy. *Health Policy*;**23**:153–66.

Gelijns A, Rosenberg N, 1994. The dynamics of technological change in medicine. *Health Aff (Millwood)*;**13**:28–46.

Gellman DD, 1993. Health technology assessment. *Can Med Assoc J*;**148**:886.

George WW, 1995. The editor's corner: medical technology and competitiveness in the world market: reinventing the environment for innovation. *J Appl Biomater*;**6**:147–52.

Gershon-Cohen J, Cooley AG, 1950. Telognosis. *Radiology*;**55**:582–7.

Gitlin JN, 1986. Teleradiology. *Radiol Clin North Am*;**24**:55–68.

Goodman C, 1992. It's time to rethink health care technology assessment. *Int J Technol Assess Health Care*;**8**:335–58.

Goodman GR, Hunter JG, 1991. Results of laparoscopic cholecystectomy in a university hospital. *Am J Surg*;**162**:576–9.

Grace PA, Qureshi A, Coleman J, 1991. Reduced postoperative hospitalization after laparoscopic cholecystectomy. *Br J Surg*;**78**:160–2.

Graffis R, 1992. Laparoscopic cholecystectomy. The Methodist Hospital experience. *Surg Laparosc Endosc*;**2**:69–73.

Graves HA, Ballinger JF, Anderson WJ, 1991. Appraisal of laparoscopic cholecystectomy. *Ann Surg*;**213**:655–64.

Griffiths V, 1995. Genetic testing under the microscope – screening has far-reaching economic and ethical implications. *Financial Times* January 3:**8**.

Grigsby J, 1995. Current status of domestic telemedicine. *J Med Syst*;**19**:19–27.

Grimes DA, 1993. Technology follies: the uncritical acceptance of medical innovation. *JAMA*;**269**:3030–3.

Gross M, 1993. Innovations in surgery: a proposal for phased clinical trials. *J Bone Joint Surg*;**75**:351–4.

Hailey DM, Crowe BL, 1991. Health technology assessment: an Australian perspective. *Med Prog Technol*;**17**:103–9.

Hailey DM, Cowley DE, Dankiw W, 1990. The impact of health technology assessment. *Community Health Stud*;**14**:223–34.

Hamerton JL, 1989. Chorionic villus sampling vs amniocentesis. *Lancet*;**i**:678.

Hardy K, Jones R, Fletcher H, 1992. Matched comparison of laparoscopic and open cholecystectomy: clinical, quality of life and economic study. In: Proceedings of the 3rd World Congress of Endoscopic Surgery; Bordeaux, France. Abstract 303.

Heckerling PS, Verp MS, 1994. A cost-effectiveness analysis of amniocentesis and chorionic villus sampling for prenatal genetic testing. *Med Care*;**32**:863–80.

Heckerling PS, Verp MS, Hadro TA, 1994. Preferences of pregnant women for amniocentesis or chorionic villus sampling for prenatal testing: comparison of patients' choices and those of a decision-analytic model. *J Clin Epidemiol*;**47**:1215–28.

Hendee WR, 1990. Technology transfer and cost constraints in medicine. *Invest Radiol*;**25**:745–7.

Hendee WR, 1991. Technology assessment in medicine: methods, status and trends. *Med Prog Technol*;**17**:69–75.

Hillman AL, 1996. Gene therapy: socioeconomic and ethical issues. A round table discussion. *Hum Gene Ther*;**7**:1139–44.

Hillner BE, Smith TJ, Desch CE, 1993. Principles of cost effectiveness analysis for the assessment of current and new therapies. *J Hematother*;**2**:501–6.

Hoare J, 1992. Tidal wave: new technology, medicine and the NHS. London: King's Fund.

Hod M, Royburt M, Neri A, Friedman S, Kaplan B, Ovadia J, 1994. Prenatal diagnosis of fetal chromosomal anomalies: current opinion on amniocentesis versus chorionic villus sampling. *Isr J Med Sci*;**30**:714–16.

Holohan TV, 1993. Assessment of new technology: Regulatory agency perspective. *J Laparoendosc Surg*;**3**:389–93.

Huston P, 1992. Is health technology assessment medicine's rising star? *Can Med Assoc J*;**147**:1839–41.

Hutton J, 1993. Assessment of medical technology: the role of engineers. *Med Biol Eng Comput*;**31**:11–15HTA.

Institute of Medicine, 1985. Assessing medical technologies. Washington DC: National Academy Press.

Jackson LG, Zachary JM, Fowler SE, *et al.,* 1992. A randomized comparison of transcervical and transabdominal chorionic-villus sampling. *N Engl J Med*;**327**:594–8.

Jacob R, Battista RN, 1993. Assessing technology assessment. Early results of the Quebec experience. *Int J Technol Assess Health Care*;**9**:564–72.

Jamison DT, Saxenian H, Bergevin Y, 1995. Investing in health wisely: the role of needs-based technology assessment. *Int J Technol Assess Health Care*;**11**:673–84.

Jan YY, Chen MF, 1993. Laparoscopic versus open cholecystectomy: a prospective randomized study. *J Formos Med Assoc*;**92**:243–9S.

Jefferson T, 1996. Economic evaluations to aid decision to conduct a trial. *Lancet*;**348**:141.

Jennett B, 1992. Health technology assessment. *BMJ*;**305**:67–8.

Johannesson M, 1994. The concept of cost in the economic evaluation of health care: a theoretical inquiry. *Int J Technol Assess Health Care*;**10**:675–82.

Johansen KS, 1988. WHO concept on health technology assessment. *Health Policy*;**9**:349–51.

Johnson LG, 1995. Gene therapy for cystic fibrosis. *Chest*;**107**:77–83S.

Johnson LG, 1996. A perspective on gene therapy for cystic fibrosis. *Expert Opin Invest Drugs*;**5**:29–42.

Jonsson B, 1993. Economic evaluation of health care technologies. *Acta Endocrinol (Copenh)*;**128** suppl 2:50–4.

Josefson D, 1996. US doctors warned about test for breast cancer gene. *BMJ*;**312**:1057–8.

Kaplan P, Normandin J Jr, Wilson GN, Plauchu H, Lippman A, Vekemans M, 1990. Malformations and minor anomalies in children whose mothers had prenatal diagnosis: comparison between CVS and amniocentesis. *Am J Med Genet*;**37**:366–70.

Kesteloot K, Penninckx F, 1993. Costs and effects of open versus laparoscopic cholecystectomies. *Health Econ*;**2**:303–12.

Kikuchi M, 1991. Status of medical engineering technology assessment in Japan. *Front Med Biol Eng*;**3**:3–15.

Knowles MR, Hohneker KW, Zhou Z, *et al.,* 1995. A controlled study of adenoviral-vector-mediated gene transfer in the nasal epithelium of patients with cystic fibrosis. *N Engl J Med*;**333**:823–31.

Koch DD, 1996. How to evaluate and implement new technologies in an era of managed care and cost containment. *Clin Chem*;**42**:797–802.

Kronborg O, 1989. Randomized controlled trials in surgery. *Theor Surg*;**4**:126–7.

Kunz R, Orth K, Vogel J, *et al.,* 1992. Laparoscopic cholecystectomy versus mini-lap-cholecystectomy. Results of a prospective, randomized study. *Chirurg*;**63**:291–5.

Kurzawinski T, Hayter B, Tate J, 1992. Cost implications of laparoscopic vs open cholecystectomy. *Gut*;**33**:6S.

Lantos J, 1994. Ethics, randomization, and technology assessment. *Cancer*;**74**(suppl 9):2653–6.

Larson GM, Vitale GC, Casey J, 1992. Multipractice analysis of laparoscopic cholecystectomy. *Am J Surg*;**163**:221–6.

Legorreta AP, Silber JH, Constantino GN, Kohyliuski RW, Zatz SL, 1993. Increased cholecystectomy rate after the introduction of laparoscopic cholecystectomy. *JAMA*;**270**:1429–42.

Lerman C, Croyle R, 1994. Psychological issues in genetic testing for breast cancer susceptibility. *Arch Intern Med*;**154**:609–16.

Lerman C, Daly M, Masny A, Balshem A, 1994. Attitudes about genetic testing for breast-ovarian cancer susceptibility. *J Clin Oncol*;**12**:843–50.

Lilford RJ, Jackson J, 1995. Equipoise and the ethics of randomization. *J R Soc Med*;**88**:552–9.

Linton AL, Naylor CD, 1990. Organized medicine and the assessment of technology. Lessons from Ontario. *N Engl J Med*;**323**:1463–7.

Lippman A, Perry TB, Mandel S, Cartier L, 1985. Chorionic villi sampling: women's attitudes. *Am J Med Genet*;**22**:395–401.

Lippman A, Tomkins DJ, Shime J, Hamerton JL, 1992. Canadian multicentre randomized clinical trial of chorion villus sampling and amniocentesis. Final report. *Prenat Diagn*;**12**:385–408.

Littell CL, 1994. Innovation in medical technology: reading the indicators. *Health Aff (Millwood)*;**13**:226–35.

Littenberg B, 1992. Technology assessment in medicine. *Acad Med*;**67**:424–8.

Liu DTY, Symonds EM, Golbus MS, 1987. Chorion villus sampling. London: Chapman and Hall.

Love JW, 1975. Drugs and operations: some important differences. *JAMA*;**232**:37–8.

Luce BR, 1995. Policy implications of modeling the cost-effectiveness of health care technologies. *Drug Inform J*;**29**:1469–75.

Luce BR, Brown RE, 1995. The use of technology assessment by hospitals, health maintenance organizations, and third-party payers in the United States. *Int J Technol Assess Health Care*;**11**:79–92.

Lumsdon K, 1992. Beyond tech assessment: balancing needs, strategy. *Hospitals*;**66**:20–6.

MacIntyre NR, 1995. Efficacy assessment criteria based on risk and cost. *Respir Care*;**40**:968–70.

Macintyre IMC, Wilson RG, 1993. Laparoscopic cholecystectomy. *Br J Surg*;**80**:552–9.

Maclean JR, 1996. Telemedicine in remote health care. Volume 1. University of Aberdeen.

Maisey MN, Lewis MK, 1994. NHS research and development: a research strategy for nuclear medicine. *Nucl Med Commun*;**15**:783–5.

Majeed AW, Troy G, Nicholl JP, *et al.,* 1996. Randomised, prospective, single-blind comparison of laparoscopic versus small-incision cholecystectomy. *Lancet*;**347**:989–94.

Marchese CA, Viora E, Campogrande ME, 1986. Cost-effectiveness ratio of chorionic villi biopsy for the prenatal diagnosis of chromosomal aberrations as compared with amniocentesis. *Ricerca Clin Lab*;**16**:533.

Martin IG, Holdsworth PJ, Asker J, *et al.,* 1992. Laparoscopic cholecystectomy as a routine procedure for gallstones: results of an 'all-comers' policy. *Br J Surg*;**79**:807–10.

Mastroiacovo P, Botto LD, Cavalcanti DP, *et al.,* 1992. Limb anomalies following chorionic villus sampling: a registry based case-control study. *Am J Med Genet*;**44**:856–64.

McCarthy M, 1996. US lab starts breast-cancer-gene screening. *Lancet*;**347**:1033.

McConnell EA, 1993. Six honest serving men. *Hosp Materiel Manage Q*;**15**:31–41.

McDonough R, 1993. Current technology assessment programs/procedures. *J Laparoendosc Surg*;**3**:395–7.

McGinn FP, Miles AJG, Uglow M, Ozmen M, Terzi C, Humby M, 1995. Randomized trial of laparoscopic cholecystectomy and minicholecystectomy. *Br J Surg*;**82**:1374–7.

McGovern MM, Goldberg JD, Desnick RJ, 1986. Acceptability of chorionic villi sampling for prenatal diagnosis. *Am J Obstet Gynecol*;**155**:25–9.

McGregor M, 1994. Can our health services be saved by technology evaluation? The Quebec experience. *Clin Invest Med*;**17**:334–42.

McIntyre RC, Zoeter MA, Weil KC, Cohen MM, 1992. Comparison of outcome and cost of open vs laparoscopic cholecystectomy. *J Laparoendosc Surg*;**2**:143–8.

McKernan JB, 1991. Laparoscopic cholecystectomy. *Am Surg*;**57**:311–12.

McLaren P, Ball CJ, 1995. Telemedicine: lessons remain unheeded. *BMJ*;**310**:1390–1.

McMahon AJ, O'Dwyer PJ, Russell IT, Baxter JN, 1992. Laparoscopic versus open cholecystectomy and the need for a randomized trial: a survey of surgeons and ethical committees in the British Isles. *J Laparoendosc Surg*;**2**:277–80.

McMahon AJ, Baxter JN, Kenny G, O'Dwyer PJ, 1993a. Ventilatory and blood gas changes during laparoscopic and open cholecystectomy. *Br J Surg*;**80**:1252–4.

McMahon AJ, O'Dwyer PJ, McMillan DC, *et al.,* 1993b. Comparison of metabolic responses to laparoscopic and minilaparotomy cholecystectomy. *Br J Surg*;**80**:1255–8.

McMahon AJ, Russell IT, Baxter JN, *et al.,* 1994a. Laparoscopic versus minilaparotomy cholecystectomy: a randomised trial. *Lancet*;**343**:135–8.

McMahon AJ, Russell IT, Ramsay G, *et al.,* 1994b. Laparoscopic and minilaparotomy cholecystectomy: a randomized trial comparing postoperative pain and pulmonary function. *Surgery*;**115**:533–9.

McMahon AJ, Ross S, Baxter JN, *et al.,* 1995. Symptomatic outcome one year after laparoscopic and minilaparotomy cholecystectomy: a randomized trial. *Br J Surg*;**82**:1378–82.

Mendelson DN, Abramson RG, Rubin RJ, 1995. State involvement in medical technology assessment. *Health Aff (Millwood)*;**14**:83–98.

Menon D, 1993. Technology assessment and biomedical engineering education. *Med Biol Eng Comput*;**31**:33–6HTA.

Menon D, Marshall D, 1996. The internationalization of health technology assessment. *Int J Technol Assess Health Care*;**12**:45–51.

Menon D, Fung F, Harstall C, O'Connell P, 1995. Development of a health technology assessment program: the case of Alberta. *Int J Technol Assess Health Care*;**11**:93–101.

Merrell RC, 1995. Telemedicine in the 90s: beyond the future. *J Med Syst*;**19**:15–18.

Michaelis AR, 1990. From the electric telegraph to the ISDN. *Telecommun J*;**57**:228–9.

Modell B, 1985. Chorionic villus sampling. Evaluating safety and efficacy. *Lancet*;**i**:737–40.

Moloney TW, Rogers DE,1979. Medical technology – a different view of the contentious debate over costs. *N Engl J Med*;**301**:1413–19.

MRC, 1994. Health technology assessment in surgery: the role of the randomized controlled trial. London: Medical Research Council.

MRC Working Party on the Evaluation of Chorion Villus Sampling, 1991. Medical Research Council European trial of chorion villus sampling. *Lancet*;**337**:1491–9.

Mugford M, Drummond MF, Henderson J, Mooney G, 1985. Chorionic villus sampling. *Lancet*;**ii**:384–5.

Muggah H, Hunter AG, Ivey B, Cox DM, 1987. Difficulties encountered in a randomization trial of CVS versus amniocentesis for prenatal diagnosis. *Clin Genet*;**32**:235–9.

Muller C, 1991. Objective health care technology evaluation – it isn't easy. *Soc Work Health Care*,**16**:119–32.

Murphy JR, 1991. Assessment process: a microscopic view. *Med Prog Technol*,**17**:77–83.

Murray TH, 1994. Assessing genetic technologies: two ethical issues. *Int J Technol Assess Health Care*,**10**:573–82.

Neugebauer E, Troidl H, Spangenberger W, Dietrich A, Lefering R, 1991. Conventional versus laparoscopic cholecystectomy and the randomised controlled trial. *Br J Surg*,**78**:150–4.

NHS Executive, 1996. Report of the NHS Health Technology Assessment Programme. London: NHS Executive.

Nightingale SL, 1993. Device 'fast track' plan. *JAMA*,**270**:693.

Nottle PD, Wale RJ, Johnson WR, 1991. Percutaneous laparoscopic cholecystectomy: the first fifty. *Aust N Z J Surg*,**61**:254–60.

Nuffield Council on Bioethics, 1993. Genetic screening: ethical issues. London: Nuffield Council on Bioethics.

Offit K, 1996. Germline BRCA1 185delAG mutations in Jewish women with breast cancer. *Lancet*,**347**:1643–5.

Olney RS, Khoury MJ, Alo CJ, *et al.,* 1995. Increased risk for transverse digital deficiency after chorionic villus sampling: results of the United States Multistate Case–Control Study, 1988–1992. *Teratology*,**51**:20–9.

Ong BN, 1996. Lay perspective in health technology assessment. *Int J Technol Assess Health Care*,**12**:511–17.

Orrell RW, Lilford RJ, 1990. Chorionic villus sampling and rare side-effects: will a randomised controlled trial detect them? *Int J Gynaecol Obstet*,**32**:29–34.

Patail BM, Aranha ANF, 1995. Role of the biomedical engineering department in William Beaumont Hospital's technology assessment process. *J Clin Eng*,**20**:290–6.

Peacock JN, Johnson AG, Squirrell D, 1995. Metabolic response to surgery and postoperative surgery following laparoscopic or mini-laparotomy cholecystectomy. Unpublished.

Pearson VAH, 1994. Minimal access surgery: a review. Bristol: Health Care Evaluation Unit, Department of Epidemiology and Public Health Medicine, University of Bristol.

Peckham M, 1991a. Research and development for the NHS. *Lancet*,**338**:367–71.

Peckham M, 1991b. Research for health: an R&D strategy for the NHS. London: HMSO.

Peckham M, 1993. Research for health. London: Department of Health.

Perednia DA, Brown NA, 1995. Teledermatology: one application of telemedicine. *Bull Med Libr Assoc*,**83**:42–7.

Perissat J, 1993. Laparoscopic cholecystectomy: the European experience. *Am J Surg*,**165**:444–9.

Perissat J, Collet D, Vitale G, Belliard R, Sosso M, 1991. Laparoscopic cholecystectomy using intracorporeal lithotripsy. *Am J Surg*,**161**:371–6.

Perry S, Hanft R, Chrzanowski R, 1993. Report from the Canadian Coordinating Office of Health Technology Assessment (CCOHTA). Satellite symposium to the eighth annual meeting of the International Society of Technology Assessment in Health Care. *Int J Technol Assess Health Care*,**9**:311–12.

Perry S, Wilkinson SL, 1992. The technology assessment and practice guidelines forum. *Int J Technol Assess Health Care*,**8**:289–300.

Peters JH, Ellison EC, Innes JT, 1991. Safety and efficacy of laparoscopic cholecystectomy. A prospective analysis of 100 initial patients. *Ann Surg*,**213**:3–12.

Phelps CE, Parente ST, 1990. Priority setting in medical technology and medical practice assessment. *Med Care*,**28**:703–23.

Pollock AV, 1989. Rise and fall of the random controlled trial in surgery. *Theor Surg*,**4**:163–70.

Ponder BA, 1995. Genetic testing of the common cancers: potential and problems. *Eur J Cancer*,**31A**:807–8.

Powe NR, Griffiths RI, 1995. Clinical-economic trial: promise, problems and challenges. *Control Clin Trials*,**16**:377–94.

Power EJ, 1995. Identifying health technologies that work. *JAMA*,**274**:205.

Power EJ, Tunis SR, Wagner JL, 1994. Technology assessment and public health. *Ann Rev Public Health*,**15**:561–79.

Purtillo DT, Paquin L, Gindhart T, 1978. Genetics of neoplasia – impact of ecogenetics on oncogenesis. *Am J Path*,**91**:609–88.

Putensen-Himmer G, Putensen C, Lammer H, Lingnau W, Aigner F, Benzer H, 1992. Comparison of postoperative respiratory function after laparoscopy or open laparotomy for cholecystectomy. *Anesthesiology*,**77**:675–80.

Rappa MA, Debackere K, 1992a. Monitoring progress in R&D communities. In: Weingart P, Sehringer R, Winter-hager M, editors. Representations of science and technology. Proceedings of the international conference on science and technology indicators; 1990 June 10–12; Bielefeld, Federal Republic of Germany. Leiden: DSWO Press.

Rappa MA, Debackere K, 1992b. Technological communities and the diffusion of knowledge. *R&D Manage*,**22**:209–20.

Rappa MA, Garud R, 1992. Modeling contribution spans of scientists in a field: the case of cochlear implants. *R&D Manage*,**22**:337–48.

Rappa MA, Debackere K, Garud R, 1992. Technological progress and the duration of contribution spans. *Technol Forecast Social Change*,**42**:133–45.

Reddick EJ, Olsen DO, 1989. Laparoscopic laser cholecystectomy: a comparison with mini-lap cholecystectomy. *Surg Endosc;***3**:131–3.

Reiser SJ, 1992. Out of chaos: a rational approach to assessing technology. *Hospitals;***66**:22–3.

Rettig RA, 1991. Technology assessment – an update. *Invest Radiol;***26**:165–73.

Rhoads GG, Jackson LG, Schlesselman SE, *et al.,* 1989. The safety and efficacy of chorionic villus sampling for early prenatal diagnosis of cytogenetic abnormalities. *N Engl J Med;***320**:609–17.

Rinde E, Nordrum I, Nymo BJ, 1993. Telemedicine in rural Norway. *World Health Forum;***14**:71–7.

Rogers E, 1983. Diffusion of innovations. New York: The Free Press.

Rona RJ, Beech R, 1993. The process of evaluation of a new technology: genetic services and the introduction of DNA probes. *J Public Health Med;***15**:185–91.

Ross G, Erickson R, Knorr D, *et al.,* 1996. Gene therapy in the United States: a five-year status report. *Hum Gene Ther;***7**:1781–90.

Royal HD, 1994. Technology assessment: scientific challenges. *AJR Am J Roentgenol;***163**:503–7.

Russell I, 1996. Can it work? Does it work? Research design for health technology assessment. York: University of York.

Rutten FFH, Bonsel GJ, 1992. High cost technology in health care: a benefit or a burden? *Soc Sci Med;***35**:567–77.

Rutten F, Drummond M, 1994. Making decisions about health technologies: a cost-effectiveness perspective. York: Centre for Health Economics, University of York.

Sainfort F, 1991. Evaluation of medical technologies: a generalized ROC analysis. *Med Decis Making;***11**:208–20.

Santarelli E, Piergiovanni R, 1996. Analyzing literature-based innovation output indicators: the Italian experience. *Res Policy;***25**:689–711.

Schauer PR, Luna J, Ghiatas AA, Glen ME, Warren JM, Sirinek KR, 1993. Pulmonary function after laparoscopic cholecystectomy. *Surgery;***114**:389–97.

Schemmer G, Johnson A, 1993. Genetic amniocentesis and chorionic villus sampling. *Obstet Gynecol Clin North Am;***20**:497–521.

Schirmer BD, Edge SB, Dix J, Hyster MJ, Hanks JB, Jones RS, 1991. Laparoscopic cholecystectomy. *Ann Surg;***213**:665–76.

Schulman KA, Yabroff KR, Glick H, 1995. A health services approach for the evaluation of innovative pharmaceutical and biotechnology products. *Drug Inform J;***29**:1405–14.

Scott WW Jr, Rosenbaum JE, Ackerman SJ, *et al.,* 1993. Subtle orthopedic fractures: teleradiology workstation versus film interpretation. *Radiology;***187**:811–15.

Sculpher M, 1993. A snip at the price? A review of the economics of minimal access surgery. Uxbridge: Brunel University.

Sculpher M, Drummond M, Buxton M, 1995. Economic evaluation in health care research and development: undertake it early and often. Uxbridge: Health Economics Research Group, Brunel University.

Shaffer MD, Shaffer MJ, 1995. Support for biomedical equipment decision making. *Biomed Instrum Technol;***29**:193–200.

Sheldon TA, Faulkner A, 1996. Vetting new technologies: those whose efficacy and safety have not been established will now be registered and evaluated. *BMJ;***313**:508.

Sheldon TA, Long A, Hopkins A, 1992. Health technology assessment. *BMJ;***305**:426.

Shulman LP, Elias S, 1993. Amniocentesis and chorionic villus sampling. *West J Med;***159**:260–8.

Sibbald WJ, Eberhard JA, Inman KJ, Sprung CL, 1993. New technologies, critical care, and economic realities. *Crit Care Med;***21**:1777–80.

Sidebottom E, 1995. Gene watch. *Bandolier;***2**:1–2.

Simoni G, Brambati B, Danesino C, *et al.,* 1983. Efficient direct chromosome analyses and enzyme determinations from chorionic villi samples in the first trimester of pregnancy. *Hum Genet;***63**:349–57.

Simpson RL, 1993. Managing innovative technology. *Nurs Manage;***24**:18–19.

Sisk JE, Glied SA, 1994. Innovation under federal health care reform. *Health Aff (Millwood);***13**:82–97.

Sloan FA, Valvona J, Perrin JM, Adamache KW, 1986. Diffusion of surgical technology: an exploratory study. *J Health Econ;***5**:31–61.

Smidt-Jensen S, Permin M, Philip J, *et al.,* 1992. Randomised comparison of amniocentesis and transabdominal and transcervical chorionic villus sampling. *Lancet;***340**:1237–44.

Smith R, 1994. Towards a knowledge based health service. *BMJ;***309**:217–18.

Smith SC, Hailey D, Drummond M, 1994. Role of economic appraisal in health technology assessment: the Australian case. *Soc Sci Med;***38**:1653–62.

Smits HL, Baum A, 1995. Health Care Financing Administration (HCFA) and reimbursement in telemedicine. *J Med Syst;***19**:139–42.

Sniderman AD, 1996. Governance of clinical trials. *Lancet;***347**:1387–8.

Solomon MJ, McLeod RS, 1993. Clinical assessment of biomedical technology. *Dis Colon Rectum;***36**:301–7.

Soper NJ, Stockman PT, Dunnegan DL, Ashley SW, 1992. Laparoscopic cholecystectomy – the new gold standard. *Arch Surg;***127**:917–23.

Southern Surgeons Club, 1991. Prospective analysis of 1518 laparoscopic cholecystectomies. *N Engl J Med*;**324**:1073–8.

Spaw AT, Reddick EJ, Olsen DO, 1991. Laparoscopic laser cholecystectomy: analysis of 500 procedures. *Surg Laparosc Endosc*;**1**:2–7.

Spencer JW, Cox DN, 1987. Emotional responses of pregnant women to chorionic villi sampling or amniocentesis. *Am J Obstet Gynecol*;**157**:1155–60.

Spencer JW, Cox DN, 1988. A comparison of chorionic villi sampling and amniocentesis: acceptability of procedure and maternal attachment to pregnancy. *Obstet Gynecol*;**72**:714–18.

Standing Group on Health Technology, 1994. 1994 report. London: Department of Health.

Steering Committee on Future Health Scenarios, 1987. Anticipating and assessing health care technology. Dordrecht: Martinus Nijhoff.

Stephenson J, 1995. Medical technology watchdog plays unique role in quality assessment. *JAMA*;**274**:999–1001.

Stevens A, Colin-Jones D, Gabbay J, 1995. Quick and clean: authoritative health technology assessment for local health care contracting. *Health Trends*;**27**:37–42.

Stiefel R, Rizkalla E, 1995. The elements of a complete product evaluation. *Biomed Instrum Technol*;**29**:482–8.

Stix G, 1996. Is genetic testing premature? *Sci Am*;**275**:73.

Stocking B, 1985. Initiative and inertia: case studies in the NHS. London: Nuffield Provincial Hospitals Trust.

Stocking B, 1988. Factors influencing the effectiveness of mechanisms to control medical technology. In: Stocking B, editor. Expensive health technologies. Oxford: Oxford University Press.

Stocking B, 1991. Factors influencing the effectiveness of mechanisms to control medical technology. London: King's Fund Centre for Health Services Development.

Stoker ME, Vose J, O'Hara P, 1992. Laparoscopic cholecystectomy: a clinical and financial analysis of 280 operations. *Arch Surg*;**127**:589–94.

Sussman JH, 1991. Financial considerations in technology assessment. *Top Health Care Financ*;**17**:30–41.

Szczepura A, 1993. Health care technology assessment (HTA) in Europe: training for the future. Reported views of 288 organisations involved in European health care. Coventry: Comett-Assess.

Szczepura A, Cooke P, 1993. Softly, softly. *Health Serv J*; (29 July):26–7.

Szczepura A, Kankaanpaa J, 1996. Assessment of health care technologies: case studies, key concepts and strategic issues. Chichester: Wiley.

Tangalos EG, 1995. Clinical trials to validate telemedicine. *J Med Syst*;**19**:281–5.

Task Force on Principles for Economic Analysis of Health Care Technology, 1995. Economic analysis of health care technology. A report on principles. *Ann Intern Med*;**123**:61–70.

Teplensky JD, 1995. Hospital adoption of medical technology: an empirical test of alternative models. *Health Serv Res*;**30**:437–65.

Terpstra OT, 1996. Laparoscopic cholecystectomy: the other side of the coin. *BMJ*;**312**:1375–6.

Thacker SB, Berkelman RL, 1986. Surveillance of medical technologies. *J Public Health Policy*;**7**:363–77.

Thibault GE, 1992. Evaluating medical technology in the 1990s. *Trans Am Clin Climatol Assoc*;**103**:255–68.

Thornbury JR, Fryback DG, 1992. Technology assessment – an American view. *Eur J Radiol*;**14**:147–56.

Touchette N, 1996. Gene therapy: not ready for prime time. *Nature Med*;**2**:7–8.

Trondsen E, Reiertsen O, Andersen OK, Kjaersgaard P, 1993. Laparoscopic and open cholecystectomy. A prospective, randomized study. *Eur J Surg*;**159**:217–21.

Tugwell P, 1995. Technology assessment: old, new and needs-based. *Int J Technol Assess Health Care*;**11**:650–62.

Tuman KJ, Ivankovich AD, 1993. High-cost, high-tech medicine: are we getting our money's worth? *J Clin Anesth*;**5**:168–77.

US Congress, Office of Technology Assessment, 1994. Identifying health technologies that work: searching for evidence. Washington DC: US Government Printing Office.

Utterback JM, 1994. Mastering the dynamics of innovation. Boston, MA: Harvard Business School Press.

van Gennip EM, Gremy F, 1993. Challenges and opportunities for technology assessment in medical informatics. Report of a MEDINFO 92 workshop. *Med Inf*;**18**:179–84.

Voyles CR, Petro AB, Meena AL, Haick AJ, Koury AM, 1991. A practical approach to laparoscopic cholecystectomy. *Am J Surg*;**161**:365–70.

Walsh NS, 1991. Laparoscopic cholecystectomy: the first six months. *J S C Med Assoc*;**87**:263–5.

Warshaw LJ, 1994. What price screening? *Hastings Cent Rep*;**24**:46.

Watson M, Murday V, Lloyd S, Ponder B, Averill D, Eeles R, 1995. Genetic testing in breast/ovarian cancer (BRCA1) families. *Lancet*;**346**:583.

Weatherall DJ, 1991. The new genetics and clinical practice. 3rd ed. Oxford: Oxford University Press.

Welsh MJ, Smith AE, 1995. Cystic fibrosis. *Sci Am*;**273**:36–43.

Werko L, Banta D, 1995. Report from the EUR-ASSESS Project. *Int J Technol Assess Health Care*;**11**:797–9.

Wheeler KE, Porter-O'Grady T, Barrell CT, 1985. Technology: a strategic factor in hospital planning. *Health Care Manage Rev*;**10**:55–63.

White JV, 1993. Registry of laparoscopic cholecystectomy and new and evolving laparoscopic techniques. *Am J Surg*;**165**:536–40.

Whitmore K, 1993. Who, what, where, when, why and how: technology assessment in a hospital setting. *Med Biol Eng Comput*;**31**:25–27HTA.

Whittenburg C, 1995. A program proposal for new technology assessment. *AORN J*;**61**:391–2.

WHO Regional Office for Europe, 1992. Risk evaluation of chorionic villus sampling: report on a meeting. Copenhagen: WHO.

Williams AL, 1993. Technology assessment in clinical laboratory science: Becoming involved in a critical process. *Clin Lab Sci*;**6**:198–200.

Wilson RG, 1995. Case study: practical tools for improving needs-based health management and technology assessment: the PHC MAP series. *Int J Technol Assess Health Care*;**11**:709–15.

Wilson P, Leese T, Morgan WP, Kelly JF, Brigg JK, 1991. Elective laparoscopic cholecystectomy for 'all-comers'. *Lancet*;**338**:795–7.

Winkler JD, Kanouse DE, Brodsley L, Brook RH, 1986. Popular press coverage of eight National Institutes of Health consensus development topics. *JAMA*;**255**:1323–7.

Wivel NA, 1994. Gene therapy: molecular medicine of the 1990s. *Int J Technol Assess Health Care*;**10**:655–63.

Wolfe BM, Gardiner BN, Leary BF, Frey CF, 1991. Endoscopic cholecystectomy: an analysis of complications. *Arch Surg*;**126**:1192–8.

Wright R, Loughrey C, 1995. Teleradiology. *BMJ*;**310**:1392–3.

Yates JRW, 1996. Medical genetics. *BMJ*;**312**:1021–5.

Zabner J, Couture LA, Gregory RJ, Graham SM, Smith AE, Welsh MJ, 1993. Adenovirus-mediated gene transfer transiently corrects the chloride transport defect in nasal epithelia of patients with cystic fibrosis. *Cell*;**75**:207–16.

Zallen DT, 1996. Public oversight is necessary if human gene therapy is to progress. *Hum Gene Ther*;**7**:795–7.

Zucker KA, Bailey RW, Gadacz TR, Imbembo AL, 1991. Laparoscopic guided cholecystectomy. *Am J Surg*;**161**:36–44.

Zucker KA, Bailey RW, Flowers J, 1992. Laparoscopic management of acute and chronic cholecystitis. *Surg Clin North Am*;**72**:1045–67.

Laparoscopic cholecystectomy

ACOST, 1993. A report on medical research and health. Advisory Council on Science and Technology. London: HMSO.

Agnifili A, Ibi I, Guadagni S, *et al.,* 1993. Perioperative pain and stress: a comparison between video laparoscopic cholecystectomy and "open" cholecystectomy. *G Chir*;**14**:344–8.

Bailey RW, Zucker KA, Flowers JL, 1991. Laparoscopic cholecystectomy. Experience with 375 consecutive patients. *Ann Surg*;**214**:531–41.

Banta HD, 1993. Minimally invasive therapy in five European countries: diffusion, effectiveness and cost-effectiveness. *Health Policy*;**23**:1–178.

Banta HD, Schersten T, Jonsson E, 1993. Implications of minimally invasive therapy. *Health Policy*;**23**:167–78.

Barkun JS, Barkun AN, Sampalis JS, *et al.,* 1992. Randomised controlled trial of laparoscopic versus mini cholecystectomy. *Lancet*;**340**:1116–19.

Barkun JS, Caro JJ, Barkun AN, Trindade E, 1995. Cost-effectiveness of laparoscopic and mini-cholecystectomy in a prospective randomized trial. *Surg Endosc*;**9**:1221–4.

Bass EB, Pitt HA, Lillemoe KD, 1993. Cost-effectiveness of laparoscopic cholecystectomy versus open cholecystectomy. *Am J Surg*;**165**:466–71.

Baxter JN, O'Dwyer PJ, 1992. Laparoscopic or minilaparotomy cholecystectomy? *BMJ*;**304**:559–60.

Begg C, Cho M, Eastwood S *et al.,* 1996. Improving the quality of reporting of randomized controlled trials: the CONSORT statement. *JAMA*;**276**:637–9.

Berci G, Sackier JM, 1991. The Los Angeles experience with laparoscopic cholecystectomy. *Am J Surg*;**161**:382–4.

Berggren U, Gordh T, Grama D, Haglund U, Rastad J, Arvidsson D, 1994. Laparoscopic versus open cholecystectomy: hospitalization, sick leave, analgesia and trauma responses. *Br J Surg*;**81**:1362–5.

Black N, 1996. Why we need observational studies to evaluate the effectiveness of health care. *BMJ*;**312**:1215–18.

Bloor K, Maynard A, 1994. Through the keyhole. *Health Serv J*;(17 November):24–6.

Border P, 1995. Minimal access ('keyhole') surgery and its implications. London: Parliamentary Office of Science and Technology.

Bouchard S, Barkun AN, Barkun JS, Joseph L, 1996. Technology assessment in laparoscopic general surgery and gastrointestinal endoscopy: science or convenience? *Gastroenterology*;**110**:915–25.

Byrne J, Timon D, Armstrong C, Horgan PG, Quill DS, 1994. Comparison of analgesic requirements and pulmonary function in 'open' versus laparoscopic cholecystectomy. *Minim Invasive Ther*;**3**:3–6.

Coelho JC, de Araujo RP, Marchesini JB, Coelho IC, de Araujo LR, 1993. Pulmonary function after cholecystectomy performed through Kocher's incision, a mini-incision, and laparoscopy. *World J Surg*;**17**:544–6.

Cuschieri A, 1989. The laparoscopic revolution – walk carefully before we run. *J R Coll Surg Edinb*;**34**:295.

Cuschieri A, 1994. Minimal access surgery: implications for the NHS. Report from a Working Group chaired by Professor Alfred Cuschieri. Edinburgh: HMSO.

Cuschieri A, 1995. Whither minimal access surgery: tribulations and expectations. *Am J Surg*;**169**:9–19.

Cuschieri A, Dubois F, Mouiel J, *et al.,* 1991. The European experience with laparoscopic cholecystectomy. *Am J Surg*;**161**:385–7.

Davis C, Arregani M, Nagan RF, *et al.,* 1992. Laparoscopic cholecystectomy: the St Vincent experience. *Surg Laparosc Endosc*;**2**:64–9.

Department of Health, 1992. Assessing the effects of health technologies: principles, practice, proposals. London: Department of Health.

Department of Health, 1995. Proposed safety and efficacy register of new interventional procedures of the Medical Royal Colleges. London: Department of Health.

Downs SH, Black NA, Devlin HB, Royston CMS, Russell RCG, 1996. Systematic review of the effectiveness and safety of laparoscopic cholecystectomy. *Ann R Coll Surg Eng*;**78**:235–323.

Dubois F, Berthelot G, Levard H, 1991. Laparoscopic cholecystectomy: historic perspective and personal experience. *Surg Laparosc Endosc*;**1**:52–7.

Ferzli G, Kloss DA, 1991. Laparoscopic cholecystectomy: 111 consecutive cases. *Am J Gastroenterol*;**86**:1176–8.

Fullarton GM, 1992. Evaluation of the cost of laparoscopic and open cholecystectomy. *Gut*;**33**:17S.

Garcia-Caballero M, Vara-Thorbeck C, 1993. The evolution of postoperative ileus after laparoscopic cholecystectomy. A comparative study with conventional cholecystectomy and sympathetic blockade treatment. *Surg Endosc*;**7**:416–19.

Gelijns AC, Fendrick AM, 1993. Dynamics of innovation in minimally invasive therapy. *Health Policy*;**23**:153–66.

Goodman GR, Hunter JG, 1991. Results of laparoscopic cholecystectomy in a university hospital. *Am J Surg*;**162**:576–9.

Grace PA, Qureshi A, Coleman J, 1991. Reduced postoperative hospitalization after laparoscopic cholecystectomy. *Br J Surg*;**78**:160–2.

Graffis R, 1992. Laparoscopic cholecystectomy. The Methodist Hospital experience. *Surg Laparosc Endosc*;**2**:69–73.

Graves HA, Ballinger JF, Anderson WJ, 1991. Appraisal of laparoscopic cholecystectomy. *Ann Surg*;**213**:655–64.

Gross M, 1993. Innovations in surgery: a proposal for phased clinical trials. *J Bone Joint Surg*;**75**:351–4.

Hardy K, Jones R, Fletcher H, 1992. Matched comparison of laparoscopic and open cholecystectomy: clinical, quality of life and economic study. In: Proceedings of the 3rd World Congress of Endoscopic Surgery; Bordeaux, France. Abstract 303.

Jan YY, Chen MF, 1993. Laparoscopic versus open cholecystectomy: a prospective randomized study. *J Formos Med Assoc*;**92**:243–9S.

Kesteloot K, Penninckx F, 1993. Costs and effects of open versus laparoscopic cholecystectomies. *Health Econ*;**2**:303–12.

Kunz R, Orth K, Vogel J, *et al.,* 1992. Laparoscopic cholecystectomy versus mini-lap-cholecystectomy. Results of a prospective, randomized study. *Chirurg*;**63**:291–5.

Kurzawinski T, Hayter B, Tate J, 1992. Cost implications of laparoscopic vs open cholecystectomy. *Gut*;**33**:6S.

Larson GM, Vitale GC, Casey J, 1992. Multipractice analysis of laparoscopic cholecystectomy. *Am J Surg*;**163**:221–6.

Legorreta AP, Silber JH, Constantino GN, Kohyliuski RW, Zatz SL, 1993. Increased cholecystectomy rate after the introduction of laparoscopic cholecystectomy. *JAMA*;**270**:1429–42.

Lilford RJ, Jackson J, 1995. Equipoise and the ethics of randomization. *J R Soc Med*;**88**:552–9.

Macintyre IMC, Wilson RG, 1993. Laparoscopic cholecystectomy. *Br J Surg*;**80**:552–9.

Majeed AW, Troy G, Nicholl JP, *et al.,* 1996. Randomised, prospective, single-blind comparison of laparoscopic versus small-incision cholecystectomy. *Lancet*;**347**:989–94.

Martin IG, Holdsworth PJ, Asker J, *et al.,* 1992. Laparoscopic cholecystectomy as a routine procedure for gallstones: results of an 'all-comers' policy. *Br J Surg*;**79**:807–10.

McGinn FP, Miles AJG, Uglow M, Ozmen M, Terzi C, Humby M, 1995. Randomized trial of laparoscopic cholecystectomy and minicholecystectomy. *Br J Surg*;**82**:1374–7.

McIntyre RC, Zoeter MA, Weil KC, Cohen MM, 1992. Comparison of outcome and cost of open vs laparoscopic cholecystectomy. *J Laparoendosc Surg*;**2**:143–8.

McKernan JB, 1991. Laparoscopic cholecystectomy. *Am Surg*;**57**:311–12.

McMahon AJ, O'Dwyer PJ, Russell IT, Baxter JN, 1992. Laparoscopic versus open cholecystectomy and the need for a randomized trial: a survey of surgeons and ethical committees in the British Isles. *J Laparoendosc Surg*;**2**:277–80.

McMahon AJ, Baxter JN, Kenny G, O'Dwyer PJ, 1993a. Ventilatory and blood gas changes during laparoscopic and open cholecystectomy. *Br J Surg*;**80**:1252–4.

McMahon AJ, O'Dwyer PJ, McMillan DC, *et al.,* 1993b. Comparison of metabolic responses to laparoscopic and minilaparotomy cholecystectomy. *Br J Surg*;**80**:1255–8.

McMahon AJ, Russell IT, Baxter JN, *et al.,* 1994a. Laparoscopic versus minilaparotomy cholecystectomy: a randomised trial. *Lancet*;**343**:135–8.

McMahon AJ, Russell IT, Ramsay G, *et al.*, 1994b. Laparoscopic and minilaparotomy cholecystectomy: a randomized trial comparing postoperative pain and pulmonary function. *Surgery*;**115**:533–9.

McMahon AJ, Ross S, Baxter JN, *et al.*, 1995. Symptomatic outcome one year after laparoscopic and minilaparotomy cholecystectomy: a randomized trial. *Br J Surg*;**82**:1378–82.

MRC, 1994. Health technology assessment in surgery: the role of the randomized controlled trial. London: Medical Research Council.

Neugebauer E, Troidl H, Spangenberger W, Dietrich A, Lefering R, 1991. Conventional versus laparoscopic cholecystectomy and the randomised controlled trial. *Br J Surg*;**78**:150–4.

NHS Executive, 1996. Report of the NHS Health Technology Assessment Programme. London: NHS Executive.

Nottle PD, Wale RJ, Johnson WR, 1991. Percutaneous laparoscopic cholecystectomy: the first fifty. *Aust N Z J Surg*;**61**:254–60.

Peacock JN, Johnson AG, Squirrell D, 1995. Metabolic response to surgery and postoperative surgery following laparoscopic or mini-laparotomy cholecystectomy. Unpublished.

Pearson VAH, 1994. Minimal access surgery: a review. Bristol: Health Care Evaluation Unit, Department of Epidemiology and Public Health Medicine, University of Bristol.

Perissat J, 1993. Laparoscopic cholecystectomy: the European experience. *Am J Surg*;**165**:444–9.

Perissat J, Collet D, Vitale G, Belliard R, Sosso M, 1991. Laparoscopic cholecystectomy using intracorporeal lithotripsy. *Am J Surg*;**161**:371–6.

Peters JH, Ellison EC, Innes JT, 1991. Safety and efficacy of laparoscopic cholecystectomy. A prospective analysis of 100 initial patients. *Ann Surg*;**213**:3–12.

Putensen-Himmer G, Putensen C, Lammer H, Lingnau W, Aigner F, Benzer H, 1992. Comparison of post-operative respiratory function after laparoscopy or open laparotomy for cholecystectomy. *Anesthesiology*;**77**:675–80.

Reddick EJ, Olsen DO, 1989. Laparoscopic laser cholecystectomy: a comparison with mini-lap cholecystectomy. *Surg Endosc*;**3**:131–3.

Schauer PR, Luna J, Ghiatas AA, Glen ME, Warren JM, Sirinek KR, 1993. Pulmonary function after laparoscopic cholecystectomy. *Surgery*;**114**:389–97.

Schirmer BD, Edge SB, Dix J, Hyster MJ, Hanks JB, Jones RS, 1991. Laparoscopic cholecystectomy. *Ann Surg*;**213**:665–76.

Sculpher M, 1993. A snip at the price? A review of the economics of minimal access surgery. Uxbridge: Brunel University.

Soper NJ, Stockman PT, Dunnegan DL, Ashley SW, 1992. Laparoscopic cholecystectomy – the new gold standard. *Arch Surg*;**127**:917–23.

Southern Surgeons Club, 1991. Prospective analysis of 1518 laparoscopic cholecystectomies. *N Engl J Med*;**324**:1073–8.

Spaw AT, Reddick EJ, Olsen DO, 1991. Laparoscopic laser cholecystectomy: analysis of 500 procedures. *Surg Laparosc Endosc*;**1**:2–7.

Stoker ME, Vose J, O'Hara P, 1992. Laparoscopic cholecystectomy: a clinical and financial analysis of 280 operations. *Arch Surg*;**127**:589–94.

Szczepura A, Kankaanpaa J, 1996. Assessment of health care technologies: case studies, key concepts and strategic issues. Chichester: Wiley.

Terpstra OT, 1996. Laparoscopic cholecystectomy: the other side of the coin. *BMJ*;**312**:1375–6.

Trondsen E, Reiertsen O, Andersen OK, Kjaersgaard P, 1993. Laparoscopic and open cholecystectomy. A prospective, randomized study. *Eur J Surg*;**159**:217–21.

Voyles CR, Petro AB, Meena AL, Haick AJ, Koury AM, 1991. A practical approach to laparoscopic cholecystectomy. *Am J Surg*;**161**:365–70.

Walsh NS, 1991. Laparoscopic cholecystectomy: the first six months. *J S C Med Assoc*;**87**:263–5.

White JV, 1993. Registry of laparoscopic cholecystectomy and new and evolving laparoscopic techniques. *Am J Surg*;**165**:536–40.

Wilson P, Leese T, Morgan WP, Kelly JF, Brigg JK, 1991. Elective laparoscopic cholecystectomy for 'all-comers'. *Lancet*;**338**:795–7.

Wolfe BM, Gardiner BN, Leary BF, Frey CF, 1991. Endoscopic cholecystectomy: an analysis of complications. *Arch Surg*;**126**:1192–8.

Zucker KA, Bailey RW, Gadacz TR, Imbembo AL, 1991. Laparoscopic guided cholecystectomy. *Am J Surg*;**161**:36–44.

Zucker KA, Bailey RW, Flowers J, 1992. Laparoscopic management of acute and chronic cholecystitis. *Surg Clin North Am*;**72**:1045–67.

CVS

Is my baby all right? Fetal diagnostic test chorionic villus sampling implicated in birth defects. *Newsweek* 1992 June 22;62–3.

Safer tests expected – a new method of prenatal diagnosis. *Financial Times* 1996 January 12;10.

Abramsky L, Rodeck CH, 1991. Women's choices for fetal chromosome analysis. *Prenat Diagn*;**11**:23–8.

Alfirevic Z, Gosden C, Neilson JP, 1995a. Chorion villus sampling compared with amniocentesis for prenatal diagnosis. Cochrane Database of Systematic Reviews.

Alfirevic Z, Gosden C, Neilson JP, 1995b. Early amniocentesis vs chorion villus sampling. Cochrane Database of Systematic Reviews.

Blakemore KJ, 1988. Prenatal diagnosis by chorionic villus sampling. *Obstet Gynecol Clin North Am*;**15**:179–213.

Brambati B, 1995. Chorionic villus sampling. *Curr Opin Obstet Gynecol*;**7**:109–16.

Brambati B, Terzian E, Tognoni G, 1991. Randomized clinical trial of transabdominal versus transcervical chorionic villus sampling methods. *Prenat Diagn*;**11**:285–93.

Brown N, Lumley J, Tickle C, Keene J, 1996. Congenital limb reduction defects. Clues from developmental biology, teratology and epidemiology. London: HMSO.

Bryce RL, Bradley MT, McCormick SM, 1989. To what extent would women prefer chorionic villus sampling to amniocentesis for prenatal diagnosis? *Paediatr Perinat Epidemiol*;**3**:137–45.

Burton BK, Schulz CJ, Burd LI, 1992. Limb anomalies associated with chorionic villus sampling. *Obstet Gynecol*;**79**:726–30.

Byrne D, Marks K, Azar G, Nicolaides KH, 1991. Randomized study of early amniocentesis versus chorionic villus sampling: a technical and cytogenetic comparison of 650 patients. *Ultrasound Obstet Gynecol*;**1**:235–40.

Canadian Collaborative CVS–Amniocentesis Clinical Trial Group, 1989. Multicentre randomised clinical trial of chorion villus sampling and amniocentesis. *Lancet*;**i**:1–6.

Chalmers TC, 1975. Randomization of the first patient. *Med Clin North Am*;**59**:1035–8.

Chalmers I, 1997. What is the prior probability of a proposed new treatment being superior to established treatments? *BMJ*;**314**:74–5.

Chitty LS, Bobrow M, 1994. Techniques in current use in prenatal diagnosis. *Int J Technol Assess Health Care*;**10**:604–27.

Evans JA, Hamerton JL, 1996. Limb defects and chorionic villus sampling. *Lancet*;**347**:484–5.

Fahy MJ, Lippman A, 1988. Prenatal diagnosis and the Canadian Collaborative Randomized Trial of chorionic villi sampling: the physician's view. *Am J Med Genet*;**31**:953–61.

Firth HV, Boyd PA, Chamberlain P, MacKenzie IZ, Lindenbaum RH, Huson SM, 1991. Severe limb abnormalities after chorion villus sampling at 56–66 days' gestation. *Lancet*;**337**:762–3.

Froster UG, Jackson L, 1996. Limb defects and chorionic villus sampling: results from an international registry, 1992–94. *Lancet*;**347**:489–94.

Hamerton JL, 1989. Chorionic villus sampling vs amniocentesis. *Lancet*;**i**:678.

Heckerling PS, Verp MS, 1994. A cost-effectiveness analysis of amniocentesis and chorionic villus sampling for prenatal genetic testing. *Med Care*;**32**:863–80.

Heckerling PS, Verp MS, Hadro TA, 1994. Preferences of pregnant women for amniocentesis or chorionic villus sampling for prenatal testing: comparison of patients' choices and those of a decision-analytic model. *J Clin Epidemiol*;**47**:1215–28.

Hod M, Royburt M, Neri A, Friedman S, Kaplan B, Ovadia J, 1994. Prenatal diagnosis of fetal chromosomal anomalies: current opinion on amniocentesis versus chorionic villus sampling. *Isr J Med Sci*;**30**:714–16.

Jackson LG, Zachary JM, Fowler SE, *et al.,* 1992. A randomized comparison of transcervical and transabdominal chorionic-villus sampling. *N Engl J Med*;**327**:594–8.

Kaplan P, Normandin J Jr, Wilson GN, Plauchu H, Lippman A, Vekemans M, 1990. Malformations and minor anomalies in children whose mothers had prenatal diagnosis: comparison between CVS and amniocentesis. *Am J Med Genet*;**37**:366–70.

Lilford RJ, Jackson J, 1995. Equipoise and the ethics of randomization. *J R Soc Med*;**88**:552–9.

Lippman A, Perry TB, Mandel S, Cartier L, 1985. Chorionic villi sampling: women's attitudes. *Am J Med Genet*;**22**:395–401.

Lippman A, Tomkins DJ, Shime J, Hamerton JL, 1992. Canadian multicentre randomized clinical trial of chorion villus sampling and amniocentesis. Final report. *Prenat Diagn*;**12**:385–408.

Liu DTY, Symonds EM, Golbus MS, 1987. Chorion villus sampling. London: Chapman and Hall.

Marchese CA, Viora E, Campogrande ME, 1986. Cost-effectiveness ratio of chorionic villi biopsy for the prenatal diagnosis of chromosomal aberrations as compared with amniocentesis. *Ricerca Clin Lab*;**16**:533.

Mastroiacovo P, Botto LD, Cavalcanti DP, *et al.,* 1992. Limb anomalies following chorionic villus sampling: a registry based case-control study. *Am J Med Genet*;**44**:856–64.

McGovern MM, Goldberg JD, Desnick RJ, 1986. Acceptability of chorionic villi sampling for prenatal diagnosis. *Am J Obstet Gynecol*;**155**:25–9.

Modell B, 1985. Chorionic villus sampling. Evaluating safety and efficacy. *Lancet*;**i**:737–40.

MRC Working Party on the Evaluation of Chorion Villus Sampling, 1991. Medical Research Council European trial of chorion villus sampling. *Lancet*;**337**:1491–9.

Mugford M, Drummond MF, Henderson J, Mooney G, 1985. Chorionic villus sampling. *Lancet*;**ii**:384–5.

Muggah H, Hunter AG, Ivey B, Cox DM, 1987. Difficulties encountered in a randomization trial of CVS versus amniocentesis for prenatal diagnosis. *Clin Genet*;**32**:235–9.

Olney RS, Khoury MJ, Alo CJ, *et al.,* 1995. Increased risk for transverse digital deficiency after chorionic villus sampling: results of the United States Multistate Case–Control Study, 1988–1992. *Teratology*;**51**:20–9.

Orrell RW, Lilford RJ, 1990. Chorionic villus sampling and rare side-effects: will a randomised controlled trial detect them? *Int J Gynaecol Obstet*;**32**:29–34.

Rhoads GG, Jackson LG, Schlesselman SE, *et al.,* 1989. The safety and efficacy of chorionic villus sampling for early prenatal diagnosis of cytogenetic abnormalities. *N Engl J Med*;**320**:609–17.

Schemmer G, Johnson A, 1993. Genetic amniocentesis and chorionic villus sampling. *Obstet Gynecol Clin North Am*;**20**:497–521.

Shulman LP, Elias S, 1993. Amniocentesis and chorionic villus sampling. *West J Med*;**159**:260–8.

Simoni G, Brambati B, Danesino C, *et al.,* 1983. Efficient direct chromosome analyses and enzyme determinations from chorionic villi samples in the first trimester of pregnancy. *Hum Genet*;**63**:349–57.

Smidt-Jensen S, Permin M, Philip J, *et al.,* 1992. Randomised comparison of amniocentesis and transabdominal and transcervical chorionic villus sampling. *Lancet*;**340**:1237–44.

Spencer JW, Cox DN, 1987. Emotional responses of pregnant women to chorionic villi sampling or amniocentesis. *Am J Obstet Gynecol*;**157**:1155–60.

Spencer JW, Cox DN, 1988. A comparison of chorionic villi sampling and amniocentesis: acceptability of procedure and maternal attachment to pregnancy. *Obstet Gynecol*;**72**:714–18.

Taylor KM, 1985. The doctor's dilemma: physician participation in randomized clinical trials. *Cancer Treat Rep*;**69**:1095–7.

Weatherall DJ, 1991. The new genetics and clinical practice. 3rd ed. Oxford: Oxford University Press.

WHO Regional Office for Europe, 1992. Risk evaluation of chorionic villus sampling: report on a meeting. Copenhagen: WHO.

Teleradiology and teledermatology

'Distance health care' is latest medicine. *Electronic Eng Times* 1996;**29**:55.

Telemedicine: big sister is watching you. *Economist* 1997;**342**(7999):51.

Telemedicine: fad or future? *Lancet* 1995;**345**:73–4.

Bashshur RL, 1995. Telemedicine effects: cost, quality, and access. *J Med Syst*;**19**:81–91.

Bashshur R, Lovett J, 1977. Assessment of telemedicine: results of the initial experience. *Aviat Space Environ Med*;**48**:65–70.

Bashshur RL, Armstrong PA, Youssef ZI, 1975. Tele-medicine: explorations in the use of telecommunications in health care. Springfield, VA: Charles C Thomas.

Bergman R, 1993. Letting telemedicine do the walking. *Hosp Health Netw*;**67**:46–8.

Coles SFS, 1995. Telemedicine: the rise of digital healthcare. London: FT Pharmaceuticals & Healthcare Publishing.

Crump WJ, Pfeil T, 1995. A telemedicine primer. An introduction to the technology and an overview of the literature. *Arch Fam Med*;**4**:796–803.

Fuchs M, 1979. Provider attitudes toward STARPAHC: a telemedicine project on the Papago Reservation. *Med Care*;**17**:59–68.

Gayler BW, Gitlin JN, Rappaport W, Skinner FL, Cerva J, 1981. Teleradiology: an evaluation of a microcomputer-based system. *Radiology*;**140**:355–60.

Gershon-Cohen J, Cooley AG, 1950. Telognosis. *Radiology*;**55**:582–7.

Gitlin JN, 1986. Teleradiology. *Radiol Clin North Am*;**24**:55–68.

Grigsby J, 1995. Current status of domestic telemedicine. *J Med Syst*;**19**:19–27.

Maclean JR, 1996. Telemedicine in remote health care. Volume 1. University of Aberdeen.

McLaren P, Ball C, J1995. Telemedicine: lessons remain unheeded. *BMJ*;**310**:1390–1.

Merrell RC, 1995. Telemedicine in the 90s: beyond the future. *J Med Syst*;**19**:15–18.

Michaelis AR, 1990. From the electric telegraph to the ISDN. *Telecommun J*;**57**:228–9.

Perednia DA, Brown NA, 1995. Teledermatology: one application of telemedicine. *Bull Med Libr Assoc*;**83**:42–7.

Rinde E, Nordrum I, Nymo BJ, 1993. Telemedicine in rural Norway. *World Health Forum*;**14**:71–7.

Scott WW Jr, Rosenbaum JE, Ackerman SJ, *et al.,* 1993. Subtle orthopedic fractures: teleradiology workstation versus film interpretation. *Radiology*;**187**:811–15.

Smits HL, Baum A, 1995. Health Care Financing Administration (HCFA) and reimbursement in telemedicine. *J Med Syst*;**19**:139–42.

Tangalos EG, 1995. Clinical trials to validate telemedicine. *J Med Syst*;**19**:281–5.

Wright R, Loughrey C, 1995. Teleradiology. *BMJ*;**310**:1392–3.

Genetic diagnosis of susceptibility to breast cancer

All our yesterdays. *Health Serv J* 1996;**106**:25.

Detectives on the case – early detection of cancer means a higher survival rate for patients. *Financial Times* 1995 January 24;13.

Genetic testing for cancer risk: research projects being funded. *Oncology* 1994;**8**:16–18.

Beardsley T, 1996. Vital data. *Sci Am*;**274**:76–81.

Brown KS, 1996. Investigators grappling with looming disease gene concerns. *Scientist*,**10**:1–9.

Bryant H, 1996. Genetic screening for breast cancer in Ashkenazi women. *Lancet*,**347**:1638.

Christensen K, 1996. Novel genetics research and technologies present major challenges for public policy. *Genet Eng News*,**16**:4–35.

Eeles R, 1996. Testing for the breast cancer predis-position gene, BRCA1. *BMJ*,**313**:572.

Evans DG, 1995. Genetic testing for cancer predisposition: need and demand. *J Med Genet*,**32**:161.

Eyre H, Sondik E, Smith RA, Kessler L, 1995. Joint meeting on the feasibility of a study of screening premenopausal women (40–49 years) for breast cancer: April 20–21, 1994. *Cancer*,**75**:1391–403.

Griffiths V, 1995. Genetic testing under the microscope – screening has far-reaching economic and ethical impli-cations. *Financial Times* January 3:**8**.

Josefson D, 1996. US doctors warned about test for breast cancer gene. *BMJ*,**312**:1057–8.

Lerman C, Croyle R, 1994. Psychological issues in genetic testing for breast cancer susceptibility. *Arch Intern Med*,**154**:609–16.

Lerman C, Daly M, Masny A, Balshem A, 1994. Attitudes about genetic testing for breast-ovarian cancer suscepti-bility. *J Clin Oncol*,**12**:843–50.

McCarthy M, 1996. US lab starts breast-cancer-gene screening. *Lancet*,**347**:1033.

Offit K, 1996. Germline BRCA1 185delAG mutations in Jewish women with breast cancer. *Lancet*,**347**:1643–5.

Ponder BA, 1995. Genetic testing of the common cancers: potential and problems. *Eur J Cancer*,**31A**:807–8.

Sidebottom E, 1995. Gene watch. *Bandolier*,**2**:1–2.

Stix G, 1996. Is genetic testing premature? *Sci Am*,**275**:73.

Warshaw LJ, 1994. What price screening? *Hastings Cent Rep*,**24**:46.

Watson M, Murday V, Lloyd S, Ponder B, Averill D, Eeles R, 1995. Genetic testing in breast/ovarian cancer (BRCA1) families. *Lancet*,**346**:583.

Yates JRW, 1996. Medical genetics. *BMJ*,**312**:1021–5.

Gene therapy for cystic fibrosis

A triumph of hype over experience? Gene therapy has promised much. It has yet to deliver the goods. *Economist* 1996;**337**(7945):111–12.

Boucher RC, Knowles MR, Johnson LG, *et al.,* 1994. Gene therapy for cystic fibrosis using E1-deleted adenovirus: a phase I trial in the nasal cavity. *Hum Gene Ther*,**5**:615–39.

Caplen NJ, Gao X, Hayes P, *et al.,* 1994. Gene therapy for cystic fibrosis in humans by liposome-mediated DNA transfer: the production of resources and the regulatory process. *Gene Ther*,**1**:139–47.

Caplen NJ, Alton EW, Middleton PG, *et al.,* 1995. Liposome-mediated CFTR gene transfer to the nasal epithelium of patients with cystic fibrosis. *Nature Med*,**1**:39–46.

Crystal RG, McElvaney NG, Rosenfeld MA, *et al.,* 1994. Administration of an adenovirus containing the human CFTR cDNA to the respiratory tract of individuals with cystic fibrosis. *Nat Genet*,**8**:42–51.

Crystal RG, Mastrangeli A, Sanders A, *et al.,* 1995a. Evalu-ation of repeat administration of a replication deficient, recombinant adenovirus containing the normal cystic fibrosis transmembrane conductance regulator cDNA to the airways of individuals with cystic fibrosis. *Hum Gene Ther*,**6**:667–703.

Crystal RG, Jaffe A, Brody S, *et al.,* 1995b. A phase 1 study, in cystic fibrosis patients, of the safety, toxicity, and bio-logical efficacy of a single administration of a replication deficient, recombinant adenovirus carrying the cDNA of the normal cystic fibrosis transmembrane conductance regulator gene in the lung. *Hum Gene Ther*,**6**:643–66.

Dorin J, 1996. Somatic gene therapy: optimism tempered by reality. *BMJ*,**312**:323–4.

European Working Group on Human Gene Transfer and Therapy Central Office, 1996. Human gene therapy clinical trials in Europe. *Hum Gene Ther*,**7**:1258–9.

Flotte TE, 1996. Clinical protocol. A phase I study of an adeno-associated virus-CFTR gene vector in adult CF patients with mild lung disease. *Hum Gene Ther*,**7**:1145–59.

Hillman AL, 1996. Gene therapy: socioeconomic and ethical issues. A round table discussion. *Hum Gene Ther*,**7**:1139–44.

Johnson LG, 1995. Gene therapy for cystic fibrosis. *Chest*,**107**:77–83S.

Johnson LG, 1996. A perspective on gene therapy for cystic fibrosis. *Expert Opin Invest Drugs*,**5**:29–42.

Knowles MR, Hohneker KW, Zhou Z, *et al.,* 1995. A controlled study of adenoviral-vector-mediated gene transfer in the nasal epithelium of patients with cystic fibrosis. *N Engl J Med*,**333**:823–31.

Ross G, Erickson R, Knorr D, *et al.,* 1996. Gene therapy in the United States: a five-year status report. *Hum Gene Ther*,**7**:1781–90.

Touchette N, 1996. Gene therapy: not ready for prime time. *Nature Med*,**2**:7–8.

Welsh MJ, Smith AE, 1995. Cystic fibrosis. *Sci Am*,**273**:36–43.

Wivel NA, 1994. Gene therapy: molecular medicine of the 1990s. *Int J Technol Assess Health Care*,**10**:655–63.

Zabner J, Couture LA, Gregory RJ, Graham SM, Smith AE, Welsh MJ, 1993. Adenovirus-mediated gene transfer transiently corrects the chloride transport defect in nasal epithelia of patients with cystic fibrosis. *Cell*,**75**:207–16.

Zallen DT, 1996. Public oversight is necessary if human gene therapy is to progress. *Hum Gene Ther*,**7**:795–7.

Appendix 1

Search strategies

Timing

Searches were made for papers on the timing of HTA using the databases, Medline and Embase. The search strategies, shown in the boxes below, were updated periodically until October 1996. Search strategies on Medline combined Medical Subject Heading (MeSH) and truncated textword terms.

Medline search strategy
1. 'TECHNOLOGY ASSESSMENT, BIOMEDICAL'/
2. HEALTH TECHNOLOG$ ASSESS$.TW
3. HEALTH CARE TECHNOLOG$ ASSESS$.TW
4. HEALTHCARE TECHNOLOG$ ASSESS$.TW
5. HEALTH TECHNOLOG$ EVALUATION$.TW
6. HEALTH CARE TECHNOLOG$ EVALUATION$.TW
7. HEALTHCARE TECHNOLOG$ EVALUATION$.TW
8. BIOMEDICAL TECHNOLOG$ EVALUATION$.TW
9. BIOMEDICAL TECHNOLOG$ ASSESS$.TW
10. 1 or 2 or 3 or 4 or 5 or 6 or 7 or 8 or 9

Embase search strategy
1. HEALTH TECHNOLOGY ASSESSMENT
2. HEALTH CARE TECHNOLOGY ASSESSMENT
3. HEALTHCARE TECHNOLOGY ASSESSMENT
4. BIOMEDICAL TECHNOLOGY ASSESSMENT
5. HEALTH TECHNOLOGY EVALUATION
6. HEALTH CARE TECHNOLOGY EVALUATION
7. HEALTHCARE TECHNOLOGY EVALUATION
8. BIOMEDICAL TECHNOLOGY EVALUATION
9. Combine 1, 2, 3, 4, 5, 6, 7, 8

Initially, attempts were made to implement a search strategy narrow enough to retrieve papers on timing of assessments alone but this proved unproductive. This led to papers being sought on HTA at a broad level, with the intention of analysing these for any information which they contained on timing. The ensuing search strategies, particularly on Medline, resulted in the retrieval of a large number of references (Medline 1966–July 1996: 1477 papers; Embase 1980–96: 260 papers). In order to ensure that the number of references remained at a manageable level, searches on Medline were not extended back further than 1991. The titles (and abstracts where available) of papers retrieved were scanned and those which were concerned with the principles of HTA were obtained and analysed for information on timing.

The database searches were supplemented with articles identified by other means. These included contact with experts in the field and hand-searching reference lists of key articles. The contents of the following journals were also monitored in 1996: *Bandolier, BMJ, Controlled Clinical Trials, Evidence Based Medicine, Health Service Journal, International Journal of Technology Assessment in Health Care, Journal of Health Services Research Policy, The Lancet, Quality in Health Care.*

Bibliometric study

Medline search strategy

CORONARY DISEASE (as an exploded MeSH term).

LAPAROSCOPIC CHOLECYSTECTOMY (both as an exploded MeSH term and also as a textword term).

CHORIONIC VILLI SAMPLING (as an exploded MeSH term) and also CHORION$ VILL$ SAMPLING (as a truncated textword term).

GENE THERAPY (both as an exploded MeSH term and also as a textword term) combined with CYSTIC FIBROSIS (both as an exploded MeSH term and also as a textword term).

GENETIC SCREENING (both as an exploded MeSH term and also as a textword term) combined with BREAST NEOPLASMS (as an exploded MeSH term) and also BREAST CANCER (as a textword term).

TELEMEDICINE (both as an exploded MeSH term and also as a textword term) combined with

DERMATOLOGY (both as an exploded MeSH term and also as a textword term) combined with TELEDERMATOLOGY (as a textword term).

TELERADIOLOGY (both as an exploded MeSH term and also as a textword term) combined with RADIOLOGY (both as an exploded MeSH term and also as a textword term) combined with TELEMEDICINE (both as an exploded MeSH term and also as a textword term).

Engineering Index search terms

LAPAROSCOPIC CHOLECYSTECTOMY
CHORION* VILL* SAMPLING
GENE THERAPY + CYSTIC FIBROSIS
GENETIC SCREENING + BREAST CANCER
TELERADIOLOGY
TELEDERMATOLOGY

PROMT search terms

LAPAROSCOPIC (W) CHOLECYSTECTOM?
(CHORION OR CHORIONIC) (W) (VILLUS OR VILLI) (W) (SAMPL?)
GENE (W) THERAPY (N) CYSTIC (W) FIBROSIS
GENETIC (W) SCREENING AND
BREAST (W) CANCER
TELERADIOLOGY
TELEDERMATOLOGY

Laparoscopic cholecystectomy

Criteria for inclusion

These included RCTs, major observational studies and systematic reviews of laparoscopic cholecystectomy, and also reviews of the literature. Since the main focus of the review was to identify factors affecting timing decisions, it was decided not to specifically assess the quality of the evaluations which were undertaken.

Methods

Searches for papers on laparoscopic cholecystectomy were made primarily on the databases, Medline and Embase. LAPAROSCOPIC CHOLECYSTECTOMY was used as the primary search term, and the search period stretched from September 1996 back until 1987, when the first laparoscopic cholecystectomy was performed.

The Medline and Embase search strategies are shown in the boxes below.

Medline search strategy
1. 'CHOLECYSTECTOMY, LAPAROSCOPIC'/
2. LAPAROSCOPIC CHOLECYSTECTOMY.TW.
3. 1 or 2
4. Limit 3 to (classical article or clinical conference or clinical trial or clinical trial, Phase I, or clinical trial, Phase II, or clinical trial, Phase III, or clinical trial, Phase IV, or controlled clinical trial or historical article or meta-analysis or multicentre study or RCT or review or review of literature or review of reported cases).

Embase search strategy
1. (LAPAROSCOPIC CHOLECYSTECTOMY + (CLINICAL TRIAL, CLINICAL TRIALS)) @ (TI, AB, KWDS)

The searches of Medline and Embase identified 496 and 194 references, respectively. The titles were scanned and papers that dealt with evaluations of the procedure or were reviews of the literature were obtained. Other databases consulted included the Cochrane Database of Systematic Reviews, the Cochrane Database of Abstracts of Reviews of Effectiveness, PROMT (national and international business journals and newspapers), System for Information on Grey Literature in Europe (SIGLE), and the Index of Scientific and Technical Proceedings.

The database searches were supplemented with articles identified by other means. These included contacting subject specialists, hand-searching reference lists from key papers and monitoring the contents of the following journals for 1996: *Bandolier, BMJ, Controlled Clinical Trials, Evidence Based Medicine, Health Service Journal, International Journal of Technology Assessment in Health Care, Journal of Health Services Research Policy, The Lancet, Quality in Health Care.*

In addition to analysing reported assessments of laparoscopic cholecystectomy, reliance was also placed on existing reviews of the literature on the procedure, such as those by Macintyre and Wilson (1993), Sculpher (1993), Cuschieri (1994), Pearson (1994) and Border (1995), and the systematic review undertaken by Downs and colleagues (1996).

The aim of the analysis was to identify whether the papers provided information on the diffusion of laparoscopic cholecystectomy, when and how

assessments of the procedure were carried out, and what factors influenced the timing of such assessments.

Chorionic villus sampling

Criteria for inclusion

These included RCTs, major observational studies and systematic reviews of CVS, and also reviews of the literature.

Methods

In order to identify relevant papers on CVS in the literature, searches were made using the on-line databases Medline and Embase. CVS, with appropriate truncation, was used as the primary search term, with the search period from September 1996 back until 1980.

The Medline and Embase search strategies are shown in the boxes below.

Medline search strategy 1
1. 'CHORIONIC VILLI SAMPLING'/
2. CHORION$ VILL$ SAMPLING.TW.
3. 1 or 2
4. Limit 3 to (classical article or clinical conference or clinical trial or clinical trial, Phase I, or clinical trial, Phase II, or clinical trial, Phase III, or clinical trial, Phase IV, or controlled clinical trial or historical article or meta-analysis or multicentre study or RCT or review or review of literature or review of reported cases.

Medline search strategy 2
1. *CHORIONIC VILLI SAMPLING/ec

Embase search strategy 1
1. (CHORION* VILL* SAMPLING) @ (TI, AB, KWDS)
2. (CLINICAL TRIAL, CLINICAL TRIALS) @ (TI, AB, KWDS)
3. 1 + 2

Embase search strategy 2
1. CHORION* VILL* SAMPLING + (ECONOMICS, HEALTH ECONOMICS)

The Medline and Embase searches identified 209 and 34 references, respectively. The titles

were scanned and papers that included evaluations of the procedure or were reviews of the literature were obtained. A search of the Cochrane Database of Systematic Reviews identified a two further papers. Other databases consulted included PROMT, SIGLE and the Index of Scientific and Technical Proceedings.

The database searches were supplemented with articles identified by other means. These included contacting subject specialists, hand-searching of reference lists from key papers, and on-going monitoring of the contents of the following journals for 1996: *Bandolier, BMJ, Controlled Clinical Trials, Evidence Based Medicine, Health Service Journal, International Journal of Technology Assessment in Health Care, Journal of Health Services Research Policy, The Lancet, Quality in Health Care.*

The papers obtained were analysed for information on the diffusion of CVS, when and how assessments of the procedure were carried out, and what factors influenced the timing of such assessments.

Telemedicine

Criteria for inclusion

These included RCTs, major observational studies, systematic reviews of teleradiology and teledermatology, and also reviews of the literature in these fields.

Methods

In order to identify relevant papers on teleradiology and teledermatology, searches were undertaken on the databases Medline (1966 to October 1996) and Embase (1980 to October 1996). The search strategies are shown in the boxes below.

Medline search strategy – Teleradiology
1. 'TELERADIOLOGY'/
2. TELERADIOLOGY.TW.
3. or 2
4. 'TELEMEDICINE'/
5. TELEMEDICINE.TW.
6. or 5
7. 'RADIOLOGY'/
8. RADIOLOGY.TW
9. 7 or 8
10. 6 and 9
11. 3 or 10

---

continued

Medline search strategy – Teledermatology
1. 'TELEMEDICINE'/
2. TELEMEDICINE.TW
3. 1 or 2
4. 'DERMATOLOGY'/
5. DERMATOLOGY.TW
6. 4 or 5
7. TELEDERMATOLOGY.TW.
8. 3 and 6
9. 7 or 8

Embase search strategy – Teleradiology
1. (TELERADIOLOGY) @ (TI, AB, KWDS)
2. (TELEMEDICINE) @ (TI, AB, KWDS)
3. (RADIOLOGY) @ (TI, AB, KWDS)
4. 2 + 3
5. 1, 4

Embase search strategy – Teledermatology
1. (TELEDERMATOLOGY) @ (TI, AB, KWDS)
2. (TELEMEDICINE) @ (TI, AB, KWDS)
3. (DERMATOLOGY) @ (TI, AB, KWDS)
4. 2 + 3
5. 1, 4

The Medline and Embase searches on teleradiology identified 158 and 100 references, respectively, and on teledermatology identified ten and five references, respectively. A search on the bibliographic database of the Telemedicine Information Exchange identified 231 references on teleradiology and 21 references on teledermatology. The Medline search identified one reference on teleradiology classed as an RCT.

Other databases consulted included the Cochrane Database of Systematic Reviews, PROMT and SIGLE. A search of Medline and the Index of Scientific and Technical Proceedings for conferences featuring teleradiology or teledermatology identified one conference including teleradiology – the Institute of Electronic and Electrical Engineers Western Canada Conference and Exhibition on Telecommunications for Health Care: Telemetry, Teleradiology and Telemedicine, 1990.

The database searches were supplemented with articles obtained by other means. These included contacting subject specialists, hand-searching of reference lists from key papers and monitoring the contents of the following journals for 1996: *Bandolier, BMJ, Controlled Clinical Trials, Evidence Based Medicine, Health Service Journal, International Journal*

of Technology Assessment in Health Care, Journal of Health Services Research Policy, Journal of Telemedicine and Telecare, The Lancet, Quality in Health Care.

The papers obtained were analysed for information on the diffusion of teleradiology and teledermatology, when and how assessments were carried out, and what factors influenced the timing of such assessments.

Cystic fibrosis gene therapy

Criteria for inclusion

These included RCTs, major observational studies and systematic reviews of gene therapy for cystic fibrosis, and also reviews of the literature.

Methods

Searches for papers on gene therapy for cystic fibrosis were made primarily on the databases Medline and Embase.

The Medline and Embase search strategies are shown in the boxes below.

Medline search strategy
1. 'GENE THERAPY'/
2. GENE THERAPY.TW
3. 1 or 2
4. 'CYSTIC FIBROSIS'/
5. CYSTIC FIBROSIS.TW
6. 4 or 5
7. 3 and 6

Embase search strategy
1. (GENE THERAPY) @ (TI, AB, KWDS)
2. (CYSTIC FIBROSIS) @ (TI, AB, KWDS)
3. 1 + 2

The Medline and Embase searches identified 248 and 355 references, respectively. The titles were scanned and papers dealing with evaluations of the procedure or reviewing the literature were obtained. Other databases consulted included the Cochrane Database of Systematic Reviews, the Cochrane Database of Abstracts of Reviews of Effectiveness, PROMT, SIGLE, and the Index of Scientific and Technical Proceedings.

The database searches were supplemented with articles identified by other means. These included

contacting subject specialists, hand-searching reference lists in key papers and monitoring the contents of the following journals for 1996: *Bandolier, BMJ, Controlled Clinical Trials, Evidence Based Medicine, Health Service Journal, Human Gene Therapy, International Journal of Technology Assessment in Health Care, Journal of Health Services Research Policy, Lancet, Quality in Health Care.*

The papers obtained were analysed for information on the diffusion of gene therapy for cystic fibrosis, when and how assessments of the procedure were carried out, and what factors influenced the timing of such assessments.

Breast cancer genetic screening

Criteria for inclusion

These included RCTs, major observational studies and systematic reviews of genetic screening for predisposition to breast cancer, and also reviews of the literature.

Methods

In order to identify relevant papers on genetic testing for breast cancer, searches were undertaken on the databases Medline (1966 to October 1996) and Embase (1980 to October 1996). The search strategies are shown in the boxes below.

Medline search strategy
1. 'GENETIC SCREENING'/
2. GENETIC SCREENING.TW
3. 1 or 2
4. 'BREAST NEOPLASMS' / BREAST CANCER.TW
5. 4 or 5
6. 3 and 6

Embase search strategy
1. (GENETIC SCREENING) @ (TI, AB, KWDS)
2. (BREAST CANCER) @ (TI, AB, KWDS)
3. 1 + 2

The Medline and Embase searches identified 81 and 47 references, respectively. The titles were scanned and papers dealing with evaluation of genetic testing for breast cancer and reviews of the literature were obtained. Other databases consulted included the Cochrane Library, PROMT, SIGLE and the Index of Scientific and Technical Proceedings.

The database searches were supplemented with articles identified by other means. These included contacting subject specialists, hand-searching the reference lists in key papers and monitoring the contents of the following journals for 1996: *Bandolier, BMJ, Controlled Clinical Trials, Evidence Based Medicine, Health Service Journal, International Journal of Technology Assessment in Health Care, Journal of Health Services Research Policy, The Lancet, Quality in Health Care.*

The papers obtained were analysed for information on the diffusion of genetic testing for predisposition to breast cancer, when and how assessments have been carried out, and what factors influenced the timing of such assessments.

Appendix 2

Conferences

These conferences were identified by searching Medline and the Index of Scientific and Technical Proceedings.

Laparoscopic cholecystectomy

1990

Annual Meeting of the Midwest Surgical Association

98th Annual Meeting of the Western Surgical Association

Symposium of Minimal Access Surgery, Laparoscopic Cholecystectomy: State of the Art

102nd Annual Meeting of the Southern Surgical Association

Annual Scientific Session of the Society of American Gastrointestinal Endoscopic Surgeons

Clinical Congress of the American College of Surgeons: Interventional Laparoscopy

Annual Meeting of the Southern California Chapter of the American College of Surgeons

1991

Consensus Conference on Bladder Calculosis, Therapeutic Strategy, Strasbourg

43rd Annual Meeting of the Surgical Section of the American Academy of Pediatrics

99th Scientific Session of the Western Surgical Association

103rd Annual Scientific Session of the Southern Surgical Association

13th Annual Meeting of the Canadian Society for Vascular Surgery/60th Annual Meeting Royal College of Physicians and Surgeons of Canada

72nd Annual Meeting of the New England Surgical Society

32nd Annual Meeting of the Society for Surgery of the Alimentary Tract

1991 Annual Meeting of the Southeastern Surgical Congress

6th Annual Gastroenterology Symposium on Highlights of Gastroenterology in The Netherlands

43rd Annual Meeting of the Southwestern Surgical Congress

111th Annual Meeting of the American Surgical Association

48th Annual Meeting of the Central Surgical Association

Conference on Lasers in Urology, Laparoscopy and General Surgery

1992

NIH Consensus Development Conference on Gallstones and Laparoscopic Cholecystectomy

100th Scientific Session of the Western Surgical Association

54th Annual Meeting of the Society of University Surgeons

35th Annual Meeting of the Midwest Surgical Association

104th Annual Meeting of the Southern Surgical Association

79th Annual Meeting of the North Pacific Surgical Association

73rd Annual Meeting of the New England Surgical Society

76th Meeting of the Surgical Research Society

NIH Consensus Development Conference on Gallstones and Laparoscopic Cholecystectomy

79th Annual Meeting of the Societe Suisse de Chirurgie

33rd Annual Meeting of the Society for Surgery of the Alimentary Tract

Annual Meeting of the Southern California Chapter of the American College of Surgeons

44th Annual Meeting of the Southwestern Surgical Congress

Orlando Meeting on Laparoscopy in Diagnosis and Therapy

49th Annual Meeting of the Central Surgical Association

63rd Annual Meeting of the Pacific Coast Surgical Association

2nd International Meeting on Horizons in Gastroenterology

Symposium on Laser Surgery: Advanced Characterisation, Therapeutics and Systems 3

1993

6th SIS-E Annual Meeting

4th Seminar on the Formation of Growths in Hepato Gastroenterology

35th World Congress of the International Society of Surgery

Annual Meeting of the Society of American Gastrointestinal Endoscopic Surgeons

14th Annual Meeting of the Surgical Infection Society

46th Annual Meeting of the Southwestern Surgical Congress

80th Annual Meeting of the Schweizerischen Gesellschaft fuer Chirurgie

Centennial Congress of the Societe Royale Belge de Chirurgie 1893–1993

36th Annual Meeting of the Midwest Surgical Association

105th Annual Scientific Session of the Southern Surgical Association

Annual Meeting of the Society of American Gastrointestinal Endoscopic Surgeons

24th Annual Meeting of the American Pediatric Surgical Association

34th Annual Meeting of the Society for Surgery of the Alimentary Tract

61st Annual Scientific Meeting of the Southeastern Surgical Congress

Annual Meeting of the Southern California Chapter of the American College of Surgeons

13th Annual Conference on Peritoneal Dialysis, San Diego

50th Annual Meeting of the Central Surgical Association

1994

2nd European Congress of the European Association for Endoscopic Surgery

111th Congress of the Deutsch Gesellschaft fuer Chirurgie – Ambivalence of Progress: Is Less More?

Annual Meeting of the Society of American Gastrointestinal Endoscopic Surgeons

Annual Meeting of the Section on Surgery of the American Academy of Pediatrics

Annual Scientific Session of the Western Surgical Association

14th World Congress of Collegium Internationale Chirurgiae Digestivae

81st Annual Meeting of the North Pacific Surgical Association

Annual Meeting of the Society of American Gastrointestinal Endoscopic Surgeons

62nd Annual Scientific Meeting and Post Graduate Course Program of the Southeastern Surgical Congress

75th Falk Symposium on Cholestatic Liver Diseases – New Strategies for Prevention and Treatment of Hepatobiliary and Cholestatic Liver Diseases

Conference on Laser Dermatology and Plastic Surgery/Conference on Dentistry/Conference on Laser Welding II

35th Annual Meeting of the Society for Surgery of the Alimentary Tract

International Meeting on Cholestasis and Related Disorders

1995

5th European Congress of Surgery (Eurosurgery 95)

International Surgical Week

Midwest Surgical Association Meeting

60th Annual Assembly of the Schweizerischen Gesellschaft fuer Gastroenterologie und Hepatologie

82nd Annual Meeting of the North Pacific Surgical Association

International Meeting on 5 Years of Laparoscopic Cholecystectomy

103rd Scientific Session of the Western Surgical Association

3rd Endoscopic Ultrasonography Belgian Meeting

Southeastern Surgical Congress 63rd Annual Scientific Meeting and Postgraduate Course Program

112th Congress of the Deutsche Gesellschaft fuer Chirurgie on Quality Assurance through Cooperation in Surgery

24th Central European Congress on Anesthesiology

Annual Meeting of the Society of American Gastrointestinal Endoscopic Surgeons

52nd Annual Meeting of the Central Surgical Association

Annual Meeting of the Society of American Gastrointestinal Endoscopic Surgeons

1st European Congress of the IHPBA

Annual Meeting of the Southern California Chapter of the American College of Surgeons

Annual Meeting of the Society of American Gastrointestinal Endoscopic Surgeons

1996

Annual Meeting of the Society of American Gastrointestinal Endoscopic Surgeons

2nd World Congress of the International Hepato Pancreato Biliary Association

XV World Congress of Collegium Internationale Chirurgiae Digestivae

Annual Scientific Meeting of the Association of Surgeons of Great Britain and Ireland

Annual Meeting of the Southern California Chapter of the American College of Surgeons

CVS

1983

11th Study Group of the Royal College of Obstetricians and Gynaecologists

1984

Joint Meeting of the Italian Society for Inborn Error Diseases, the Italian Society of Medical Genetic, and the Italian Society of Medical Cytogenetics

International Symposium on First Trimester Fetal Diagnosis

52nd Annual Meeting of the Central Association of Obstetricians and Gynecologists

40th Annual Meeting of the Society of Obstetricians and Gynaecologists of Canada

1985

35th Annual Symposium on the Biology of the Skin

2nd International Symposium on the Fetus as a Patient – Diagnosis and Treatment

International Symposium on First Trimester Fetal Diagnosis

52nd Annual Meeting of the Pacific Coast Obstetrical and Gynecological Society

53rd Annual Meeting of the Central Association of Obstetricians and Gynecologists

5th Annual Meeting of the Society of Perinatal Obstetricians

1986

42nd Assembly of Deutsche Gesellschaft fuer Gynakologie und Geburtshilfe

92nd Ross Conference on Pediatric Research: Frontiers in Genetic Medicine

Annual Meeting of the Society for the Study of Inborn Errors in Metabolism

Annual Meeting of the Schweizerischen Gesellschaft fuer Gynakologie: Operative and Perioperative Problems at the Time of Gynaecological and Obstetrical Surgery

1987

43rd Annual Meeting of the Society of Obstetricians and Gynaecologists

Annual Meeting of the Schweizerischen Gesellschaft fuer Gynakologie

49th Annual Meeting of the South Atlantic Association of Obstetricians and Gynecologists

1988

12th World Congress of Gynecology and Obstetrics

4th International Symposium on the Fetus as a Patient

7th Annual Meeting of the Society of Obstetricians and Gynaecologists of Canada

1989

National Education Conference on Strategies in Genetic Counselling: Reproductive Genetics and New Technologies

Symposium in Perinatal Medicine – Advances in Human Genetics: Current Applications and Prospects for the Future

4th International Workshop on the Fragile X and X-Linked Mental Retardation (FRA(X) and XLMR)

51st Annual Meeting of the South Atlantic Association of Obstetricians and Gynecologists

1990

48th Congress of the Deutschen Gesellschaft fuer Gynakologie und Geburtshilfe

26th Nestle Workshop on Perinatalogy

International Conference on Placenta: Basic Science and Clinical Application for the Next Decades

5th International Congress on Early Fetal Diagnosis: Recent Progress and Public Health

7th World Congress on Human Reproduction

1991

13th World Congress of Gynaecology and Obstetrics

60th Annual Meeting of the Central Association of Obstetricians and Gynecologists

38th Annual Meeting of the Society for Gynecologic Investigation

11th Annual Meeting of the Society of Perinatal Obstetricians

53rd Annual Meeting of the South Atlantic Association of Obstetricians and Gynecologists

1992

World Health Organisation/European Regional Office (WHO/EURO) Meeting: Statement on the Use of CVS in Prenatal Diagnosis

49th Meeting of the Deutschen Gesellschaft fuer Gynakologie und Geburtshilfe

4th International Conference on Limb Development and Regeneration

International Conference on Improving Birth Quality and Child Upbringing

1993

13th Annual Meeting of the Society of Perinatal Obstetricians

1994

Annual Meeting of the Osterreichischen Gesellschaft fuer Gynakologie und Geburtshilfe

14th Annual Meeting of the Society of Perinatal Obstetricians

1995

Xth Congress of the European Association of Gynaecologists and Obstetricians

Annual Meeting of the Swiss Society for Obstetrics and Gynecology/Swiss Society for Senology

15th Annual Meeting of the Society of Perinatal Obstetricians

Telemedicine

1980

International Conference on Systems Science in Health Care

1984

International Teleconference Symposium

5th International Congress on Medical Informatics in Europe

1985

6th European Congress on Medical Informatics

1987

1987 Conference on Biomedical Technologies: Montech 87

1989

International Conference on Communications: World Prosperity through Communications (BOSTONICC/89)

International Conference on Olympus Utilisation

1990

IEEE Western Canada Conference and Exhibition on Telecommunication for Health Care: Telemetry, Teleradiology and Telemedicine (IEEE Wescanex 90)

Working Conference on Telematics in Medicine

International Symposium, Whither Computers in Diabetes Care?

International Conference on Integrated Broadband Services and Networks

1991

International Telemedicine Conference

4th International Conference on Human–Computer Interaction (HCI International)

Annual International Conference of the IEEE Engineering in Medicine and Biology Society

1992

1st European Symposium on Telepathology

14th Annual International Conference of the IEEE Engineering in Medicine and Biology Society

1993

Telemedicine and Access to Care

International Telemedicine Conference

IFIP TC6/WG6.1 International Conference on Open Distributed Processing

World Summit on Medical Education

Conference on Video Communications and PACS for Medical Applications

Multimedia Communications 1993 Conference – Forging the Link: Market Technology Policy

11th International Congress of the European Federation for Medical Informatics – MIE93

1994

2nd NASA/USUHS International Conference on Telemedicine for Remote Health Care and Disaster Response

TeleMed 94

Conference on Health in the New Communications Age – Health Care Telematics for the 21st Century

16th Annual International Conference of the IEEE Engineering in Medicine and Biology Society on Engineering Advances: New Opportunities for Biomedical Engineers

18th Annual Symposium on Computer Applications in Medical Care – Transforming Information, Changing Health Care

International Congress for Lung Cancer

Conference on Applications of Digital Image Processing XVII

IFIP TC6 International Conference on Information Networks and Data Communication

1995

TeleMed 95

Round Table Meeting on Telemedicine – Risks and Opportunities

Conference on Health Care Information Infrastructure

22nd Annual Scientific Meeting of Computers in Cardiology

Conference on Health Care Technology Policy II – The Role of Technology in the Cost of Health Care: Providing the Solutions

Conference on Emerging High Speed Local Area Networks and Wide Area Networks

International Symposium on Computer and Communication Systems for Image Guided Diagnosis and Therapy (CAR 95)

Summer Workshop on Computational Modeling and Imaging in Biosciences

IS & TS 48th Annual Conference on Imaging on the Information Superhighway

7th International Congress on Medical Librarianship – Health Information for the Global Village

11th International Symposium on the Creation of Electronic Health Record Systems and Global Congress on Patient Cards

Conference on PACS Design and Evaluation – Engineering and Clinical Issues

Proceedings of the Society of Photo-Optical Instrumentation Engineers (SPIE) Medical Imaging

3rd Medicine Meets Virtual Reality Conference (MMVRIII)

Medical Imaging 1995: Image Display Conference

Conference on Advanced Image and Video Communications and Storage Technologies

1996

TeleMed 96

Medical Informatics Europe Congress 96 (MIE 96)

3rd International Symposium on Interworking (Interworking 96)

Summer Workshop on Computational Modelling, Imaging and Visualization in Biosciences (Combio 96)

International Symposium on Computer and Communication Systems for Image Guided Diagnosis and Therapy (CAR 96)

1996 Medical Imaging Symposium on PACS Design and Evaluation – Engineering and Clinical Issues

Conference on Image Display

Healthcare Computing Conference (HC96) – Current Perspectives in Healthcare Computing

Medicine Meets Virtual Reality 4 Conference (MMVR4)

18th Annual Pacific Telecommunications Conference (PTC96)

Genetic testing/screening

1981

5th Arnold O Beckman Conference in Clinical Chemistry

6th International Congress of Human Genetics

1983

Conference on Inborn Errors of Metabolism, North Shore University Hospital, Manhasset

1984

Conference on Medical Screening and Biological Monitoring for the Effects of Exposure in the Workplace

1985

Albany Birth Defects Symposium

14th Study Group of the Royal College of Obstetricians and Gynaecologists: Litigation and Obstetrics and Gynaecology

1986

92nd Ross Conference on Pediatric Research: Frontiers in Genetic Medicine

1987

International Symposium on Familial Hypercholesterolemia

1989

1st International Conference on Genetic Variation and Nutrition

4th International Workshop on the Fragile X and X-Linked Mental Retardation

6th World Congress of In Vitro Fertilization and Alternate Assisted Reproduction

Meeting on Genetics and Biology of Alcoholism, Cold Spring Harbor

Workshop on Genetic Screening: from Newborns to DNA Typing

Symposium on Recent Advances in Hemophilia Care

1990

1st European Congress on Medullary Thyroid Carcinoma

24th Conference on Genetics, Ethics and Human Values: Human Genome Mapping, Genetic Screening and Gene Therapy

Symposium on New Technologies for Genetic and Newborn Screening

1991

Workshop on Reproductive Genetic Testing: Impact upon Women

1992

1992 International Fragile X Confernce

International Symposium on Retinal Degeneration

7th Annual San Diego Conference on Genetic Recognition

1993

2nd World Congress of Perinatal Medicine

19th National Meeting of the Clinical Ligand Assay Society

Sixth International Workshop on the Fragile X and X-Linked Mental Retardation

9th International Neonatal Screening Symposium/2nd Meeting of the International Society for Neonatal Screening

23rd European Symposium on Calcified Tissue

1994

2nd Annual Meeting of the Australian Electrophoresis Society

Conference on Beryllium-related Diseases

Conference on Technoscience and Cyberculture

59th Cold Spring Harbor symposium on Quantitative Biology – the Molecular Genetics of Cancer

1995

2nd International Research Conference on Familial Cancer

2nd Annual Meeting of the Australian Electrophoresis Society

Symposium on Air Toxics – Biomarkers in Environmental Applications

33rd Annual Symposium of the SSIEM on Lactic Acidosis and Inborn Errors of Metabolism

15th World Congress on Fertility and Sterility

Workshop on Genetic Screening for Colorectal Cancer

112th Congress of the Deutsche Gesellschaft fuer Chirurgie

56th Annual Meeting of the Society of University Surgeons

1996

Molecular Aspects in the Pathogenesis and Diagnostics of Thyroid Diseases Conference of the Thyroid Gland Section of the Deutsche Gesellschaft fuer Endokrinologie

Gene therapy

1988

Symposium on Gene Transfer and Gene Therapy

1990

24th Conference on Genetics, Ethics and Human Values: Human Genome Mapping, Genetic Screening and Gene Therapy

British Pharmaceutical Conference: Impact of the New Biologies on the Medical and Pharmaceutical Sciences

1991

Conference on Gene Therapy

1992

From Genetics to Gene Therapy Meeting, University College, London

Symposium on Foetal and Neonatal Cell Transplantation and Retroviral Gene Therapy

6th Annual North American Conference on Cystic Fibrosis

1993

Conference on Gene Therapy for Neoplastic Diseases

American Lung Association/American Thoracic Society International Conference

18th European Cystic Fibrosis Conference

1994

1st Seminar on Gene Therapy from the Antoine-Lacassagne Centre

32nd Annual Symposium of the SSIEM

1995

International Workshop on the Development and Applications of Vaccines and Gene Therapy in AIDS

Symposium on Gene Therapy – Current Status and Future Prospects

Symposium on Gene Therapy – New Frontiers

OECD Ottawa 95 Workshop on Gene Delivery Systems

XL Scientific Reunion of Argentine Society of Clinical Investigation

Appendix 3

Guidelines and on-going programmes

Practical considerations for safe MAS (Cuschieri, 1994)

1. MAS must be practised by surgeons within their respective speciality.
2. The training process involves attendance at skills centres or approved practical courses, to gain exposure to the basic techniques of endoscopic surgery.
3. Before embarking on laparoscopic surgery on his/her own, a surgeon must become proficient in diagnostic laparoscopy, should have assisted in endoscopic operations (as camera person and first assistant) and visited centres where these procedures are in routine use.
4. If at all possible, the surgeon should be proctored by an experienced colleague for the first few cases.
5. In the current situation and stage of development of MAS, no surgeon should attempt on his/her own an endoscopic operation that the surgeon has not previously performed by the open conventional approach.
6. Nursing and technical support teams should also be suitably trained in the practice of MAS, and be familiar with equipment and its handling.
7. Ideally, the setting-up of MAS should be procedure orientated. Within this framework, experience in the procedure selected (e.g. cholecystectomy) should be obtained by the team and validated by an on-going audit survey before the team progresses to other endoscopic operations.
8. Self-training is insufficient and must be supplemented by attendance at specialised con-tinuing education courses and visits to centres where procedures are established. Proctoring for new operations is actively encouraged and should be supported financially by health authorities and NHS trusts.
9. A surgeon should inform the patient that he/she is performing an operation for the first time using endoscopy, whenever this situation arises.

On-going teleradiology and teledermatology programmes

A total of 45 programmes which featured teleradiology and six which featured teledermatology were identified through the Telemedicine Information Exchange. The preponderance of American programmes reflects the strong bias of the Telemedicine Information Exchange towards the USA.

Teleradiology

Australia
Foetal Health Decision Support System and Remote Use of Ultrasound Imaging

Chile
Catholic University of Chile Telemedicine Project

Thailand
Thailand Telemedicine Project

USA
Alaska Telemedicine Project

University of California, Davis School of Medicine and Medical Center

University of Colorado Health Sciences Center Telemedicine/Distance Education Program

Colorado Telehealth Network

University of Colorado Hospital Telemedicine Program

Yale Telemedicine Center, Connecticut

Kootenai Medical Center, Idaho

Kirby Hospital, Illinois

Midwest Rural Telemedicine Consortium (MRTC), Iowa

National Laboratory for the Study of Rural Telemedicine, Iowa

KAWNET, Community Hospital of Onaga Inc., Kansas

Kentucky Telecare

Massachusetts General Hospital Telemedicine Center

WorldCare, Massachussets

Upper Peninsula Telemedicine Project, Michigan

University of Missouri School of Medicine

Washington University School of Medicine, Missouri

Montana Deaconess Medical Center

Nevada Rural Hospital Project Teleradiology Network

Bassett Healthcare Medical Telecommunications, New York

Wake Forest University Medical Centre, North Carolina

East Carolina School of Medicine

North Carolina Health Care Information and Communications Alliance, Inc.

NTIA Rural ED Telemedicine Link

Rural Eastern Carolina Health Network

West River Regional Medical Centre, North Dakota

Cleveland Clinic Foundation, Ohio

Oklahoma Medical Information Network

Konawa Community Health Center Mobile Clinic/Telemedicine, Oklahoma

Comanche County Memorial Hospital, Oklahoma

Northeastern Oregon Teleradiology Network

Pennsylvania Healthnet

Allegheny-Singer Research Institute, Pennsylvania

Department of Neurosurgical Surgery, Pittsburgh, Pennsylvania

South Dakota Telemedicine Project

Telehealth Project, Texas Children's Hospital

University of Utah Telemedicine Outreach Program

Yakima Valley Radiology, Washington

Project Seahawk, Washington

Walter Reed Army Medical Centre, Washington DC

Georgetown University Medical Center, Washington DC

Wyoming Teleradiology Consortium

Teledermatology
Finland
Teledermatology Pioneer Trial

USA
Teledermatology Outpatient Prison Clinic, Connecticut

Baltimore VA Teledermatology Pilot, Maryland

Massachussets General Hospital Telemedicine Center

Practical Teledermatology, Minnesota

Advanced Telemedicine Research Group, Oregon

Appendix 4
Analysis of studies included

This appendix contains a list of the papers on the general principles of HTA that were identified using the search strategy on timing presented in Appendix 1. Each paper listed is analysed in order to ascertain whether timing of assessment and method of assessment are discussed, and also for the general focus of the paper. The papers listed here are not all cited in the main text but are included in the list of references.

Author	Title	Type[*]	Timing discussed?	How to assess discussed?	Focus of paper
Anon. 1993	Surgical innovation under scrutiny	J	No	Yes, p187	Health technology is needed for new surgical procedures. Example of laparoscopic cholecystectomy
Abrams & Hessel, 1987	Health technology assessment: problems and challenges	J	No	No	Council on Health Care Technology Assessment of the Institute of Medicine
Adams, 1992	Economic analysis in randomized controlled trials	J	No	Yes, p231–7	Assesses the prevalence and completeness of economic analyses in RCTs published from January 1966 to June 1988
Adang *et al,* 1995	Medical technology assessment: economic evaluation of new technologies	J	No	Yes, p563–6	Describes various techniques by which an economic evaluation can be performed
ACOST, 1993	Report on medical research and health	R	Yes, p16,17	Yes, p17,29–30	To identify how advances in science and technology can be used to provide better health in the most cost-effective way
Alder, 1988	Organization key to technology diffusion	J	No	Yes, p82–3	Steps for healthcare executives to take in evaluating new technologies
Andersson (1995)	Why is the pharmaceutical industry investing increasing amounts in health economic evaluations?	J	No	No	Reasons why the pharmaceutical industry is financing an increasing number of health economic evaluations
Andrade, 1991	What, me worry? Bioengineering and the costs of health care	J	No	No	Academic bioengineering and the costs of health-care
Antczak-Bouckoms *et al,* 1991	Using medical registries and data sets for technology assessment	J	No	Yes, p123–8	Using registries and data sets as sources of data for technology assessment
Ball, 1995	Can we afford advanced and sophisticated technology in tomorrow's health care arena?	J (letter)	No	No	Financing of healthcare delivery, advanced technology and procurement
Banta, 1986	Dutch committee assesses the future of health technology	J	Yes, p19	Yes, p18	Steering Committee on Future Health Scenarios, policy and HTA
Banta, 1990	The regulation of medical devices	J	No	Yes, p697–8	Medical devices industry and regulation of medical devices
Banta, 1994	Health care technology and its assessment in eight countries	J (special issue)	Yes, chap 1, p10	Yes, chap 1–10	HTA in Australia, Canada, France, Germany, The Netherlands, UK, USA, Sweden
Banta & Andreasen, 1990	Political dimension in health care technology assessment programs	J	Yes, p117	Yes, p117	Political considerations in healthcare technology assessment

[*] *B, book; C, conference proceedings; DP, discussion paper; J, journal article; P, paper; R, report.*

Author	Title	Type*	Timing discussed?	How to assess discussed?	Focus of paper
Banta & Gelijns, 1994	Future and health care technology: implications of a system for early identification	J	Yes, p143	Yes, p143	Analysis of future and emerging healthcare technology and the potential usefulness of such an activity in a healthcare policy-making and decision-making context
Banta & Luce, 1993	Health care technology and its assessment: an international perspective	B	Yes, section III	Yes, section III	Development and diffusion; assessing healthcare technology – efficacy, safety, financial costs, quality of life, social implications; assessing prevention, medical imaging, surgical practice, drugs, PACS; international perspective
Banta & Thacker (1990)	Case for reassessment of health care technology. Once is not enough	J	Yes, p236-7, 239	Yes, p236	HTA should be an iterative process in terms of the life-cycle of a technology
Banta & van Beekum, 1990	Regulation of medical devices and quality of medical care	J	No	No	Regulation of medical devices; comparison with drugs; medical devices industry; steps to improve device quality
Barondess, 1994	Rational use of new technologies	J (letter)	No	No	Need to inculcate rigorous clinical thinking in the application of new technologies
Battista & Hodge, 1995	Development of health care technology assessment. An international perspective	J	No	Yes, p268	Nature of HTA; assessments in Canada, France, The Netherlands, UK, USA; its determinants
Battista et al, (1995)	Evaluation of the Canadian Coordinating Office for Health Technology Assessment	J	Yes, p106–7, 113	No	Evaluation of CCOHTA; key features of HTA; early warning system
Berg, 1992	Assessing a technology: what constitutes enough?	J	No	Yes, p439	Too many research programmes finish after demonstrating biological efficacy, without showing that whatever it is works in practice
Bergevin & Tugwell (1995)	Introduction: needs-based technology assessment. Who can afford not to use it?	J	No	No	Outline of needs-based technology assessment
Black, 1996	Why we need observational studies to evaluate the effectiveness of health care	J	No	Yes, p1215–18	Limitations of RCTs – they may be unnecessary, inappropriate, impossible or inadequate; observational methods and RCTs should be seen as complementary
Blais, 1991	Using administrative data bases for technology assessment in health care	J	No	Yes, p206	Potential of using administrative databases for technology assessment in health-care has not yet been fully utilised
Blanpain, 1983	Health technology assessment in Belgium, France, Germany and Japan	J	No	No	Health technology assessment in Belgium, France, Germany and Japan; financial incentives; review of technologies – drugs, devices, organisational and supportive systems
Bond, 1994	What are nursing's priorities?	J	No	No	HTA and its relevance to nursing
Briggs et al, 1994	Uncertainty in the economic evaluation of health care technologies: the role of sensitivity analysis	J	No	Yes, p99–102	Four broad areas of uncertainty – variability in sample data, generalisability of results, extrapolation of results, analytical methods employed; four types of sensitivity analysis – simple, threshold, extreme and probabilistic
Bunker, 1981	Health care technology assessment essential to effective medical care	J	No	No	Reimbursement, regulation and HTA
Bunker et al, 1978	Surgical innovation and its evaluation	J	Yes, p937, 940–1	Yes, p937–41	Early clinical trials hasten the prompt evaluation of new operations
Burke, 1994	High-tech and health reform: the folly of doing it	J	No	No	Four steps for managing medical technology in the context of a never-ending supply of new treatments

* B, book; C, conference proceedings; DP, discussion paper; J, journal article; P, paper; R, report.

Author	Title	Type*	Timing discussed?	How to assess discussed?	Focus of paper
Buto, 1994	How can Medicare keep pace with cutting-edge technology?	J	Yes, p139	Yes, p139	Potential solutions for Medicare to improve coverage on promising crossover therapies
Challah & Mays, 1986	Randomised controlled trial in the evaluation of new technology: a case study	J	No	Yes, p877–8	Barriers which exist to proper evaluation of new techniques
Chalmers, 1975	Randomization of the first patient	J	No	Yes, p1035–8	Argues that the first sick patient to receive a new drug, procedure or operation should be randomised
Chen, 1993	Biotechnology transfer at the National Institutes of Health	J	No	No	Laws, regulations and guidelines play a crucial role in the success of technology transfer
Cram *et al*, 1995	Ethical issues of life-sustaining technology	J	No	Yes, p27	Ethics of access, use, cost of life-sustaining technology; potential solutions
Crepea, 1995	A systems engineering approach to technology assessment	J	No	Yes, p297–303	Six-step technology assessment process based on systems engineering; applicable to biomedical instrumentation
Culyer, 1988	Technology assessment in Europe: its present and future roles	B (chap)	No	Yes, p54–78	Current state of the art of economic appraisal
David, 1989	Technology related decision-making issues in hospitals	C	No	No	Clinical engineering input to hospital-based technology assessment programmes
David, 1993	Technology evaluation in a US hospital: the role of clinical engineering	J	No	No	Clinical engineering input to hospital-based technology management programmes
David *et al*, 1993	New approaches to technology assessment: opportunities and trends	J	No	No	Technology assessment requires a multidisciplinary approach combining clinical, technical and financial information
Davidoff & Powe, 1996	Role of perspective in defining economic measures for the evaluation of medical technology	J	No	No	Importance of taking account of perspectives of different parties when designing economic analyses
Davies *et al*, 1994	Current status of economic appraisal of health technology in the European Community: report of the Network	Journal article	No	Yes, p1603	Report of a survey of economic evaluations in EU countries to identify the impact of the results on healthcare decision- and policy-making
de Charro, 1990	Economics of technological change in health. The conditions governing the production and diffusion of innovations in health care	J	No	No	Conditions governing the production and diffusion of healthcare innovations
Deber, 1992	Translating technology assessment into policy. Conceptual issues and tough choices	J	Yes, p136	No	Technology assessment rarely provides clear yes/no answers; more usually the question is who should receive an intervention and in what circumstances
Department of Health, 1992	Assessing the effects of health technologies: principles, practice proposals	R	Yes, p9,13,16	Yes, p5,14–17	Evaluating the effects of health technologies – outcomes, research designs; using evidence about the effects of health technologies; fostering proper assessment of the effects of health technologies
Department of Health, 1995	Proposed safety and efficacy register of new interventional procedures of the medical Royal Colleges	P	Yes, p1	Yes, p1–6	Outlines a system for ensuring the safety and efficacy of new interventional procedures
Diamond & Denton, 1993	Alternative perspectives on the biased foundations of medical technology assessment	J	No	No	Definition of several biased perspectives about technology assessment that derive from the distinction between individuals and groups

* *B, book; C, conference proceedings; DP, discussion paper; J, journal article; P, paper; R, report.*

Author	Title	Type[*]	Timing discussed?	How to assess discussed?	Focus of paper
Dickersin & Herxheimer, 1996	Introduction: the quality of the medical evidence: is it good enough?	J	No	Yes, p188	Shortcomings of research studies
Dolan & Zingg, 1993	Health care technology: how can we tell if we can afford it? A Canadian viewpoint	J	No	Yes, p280–1	Health technology assessment must include the science behind the technology, the technology itself, the devices resulting from the technology, the medical outcome and the societal impact in addition to pure cost considerations
Donaldson & Sox, 1992	Setting priorities for health technology assessment: a model process	B	No	Yes, p23–5	Setting priorities for health technology assessments – general principles, a proposed process, how to implement it
Drummond, 1990	Allocating resources	J	Yes, p88–90	Yes, p77–90	Reviews methods of economic evaluation of health technology
Drummond, 1992	Test drive	J	No	Yes, p26–7	Methods of assessment for safety, efficacy, effectiveness and efficiency of health technologies; questions that managers should ask when the application of new technology is being proposed
Drummond, 1994	Evaluation of health technology: economic issues for health policy and policy issues for economic appraisal	J	Yes, p1598	Yes, p1597–8	Considers what policy issues are amenable to economic analysis, or could be greatly informed by economic appraisal results
Drummond et al, 1992	Issues in the cross-national assessment of health technology	J	No	Yes, p673–5	How cross-national differences affect the cost-effectiveness of health technologies or their evaluation
Durand-Zaleski & Jolly, 1990	Technology assessment in health care – decision makers and health care providers: what they need to know	J	Yes, p43	Yes, p38–43	Assessing technical and medical, ethical and legal, and economic aspects of emerging health technologies
Eddy, 1990	Should we change the rules for evaluating medical technologies?	B (chap)	No	Yes, p117–34	Argues that HTA should draw on the strengths of RCTs and observational approaches to speed the acceptance and diffusion of technologies that are worth the costs and deserve priority
Elliott & Hollins, 1995	Product evaluation: theoretical and practical considerations	J	No	Yes, p16	Theoretical and practical considerations of the evaluation process
Feeny & Torrance, 1989	Incorporating utility-based quality-of-life assessment measures in clinical trials	J	No	Yes, p197–202	Incorporates utility-based quality-of-life assessment measures in clinical trials; example of Canadian CVS trial
Feeny et al, 1986	Health care technology: effectiveness, efficiency and public policy	B	Yes, chap 1	Yes, chap 4–7, 9–10	Healthcare technology and its evaluation
Ferguson et al, 1993	Court-ordered reimbursement for unproven medical technology: circumventing technology assessment	J	No	No	US cases where courts of law ruled against insurance carriers who had been sued for reimbursement for unproven medical procedures
Fineberg, 1985	Technology assessment: motivation, capability and future directions	J	Yes, p669–70	Yes, p668,670	Features current status and future directions of technology assessment; eight dimensions of analysis
Firshein, 1986	HCFA urged to speed approval of technologies	J	No	No	Health Care Administration Financing focus on containing Medicare costs is hampering the use and manufacture of new medical technology

[*] *B, book; C, conference proceedings; DP, discussion paper; J, journal article; P, paper; R, report.*

Author	Title	Type[*]	Timing discussed?	How to assess discussed?	Focus of paper
Franklin, 1993	Basic concepts and fundamental issues in technology assessment	J	No	Yes, p119	HTA – what it is, safety, effectiveness; questions to ask in evaluating effectiveness; ethical questions
Franklin, 1994	Effectiveness and efficacy – which is which?	J (letter)	No	No	Definition of efficacy and effectiveness
Fuchs & Garber, 1990	The new technology assessment	J	Yes, p675	Yes, p673–7	Technology assessment now encompasses effectiveness, quality of life, patient preferences and evaluation of costs and benefits
Gafni & Birch, 1993	Guidelines for the adoption of new technologies: a prescription for uncontrolled growth in expenditures and how to avoid the problem	J	No	Yes, p915–16	Use of Healthy Years Equivalent (HYE) or Willingness to Pay (WTP) as measures of outcome
Garber, 1994	Can technology assessment control health spending?	J	No	Yes, p118–21	Approaches to technology assessment and the ways in which technology assessment, in the form of cost-effectiveness analysis, can be applied to help control spending growth
Gelijns, 1990	Modern methods of clinical investigation	B	Yes, p173	Yes, chap 1, 3, 10	Innovation–evaluation nexus; endpoints; clinical trials; evaluating medical technologies; comparing development of drugs, devices and clinical procedures
Gelijns & Rosenberg, 1994	The dynamics of technological change in medicine	J	Yes, p44	Yes, p44	Contrasts a dynamic and interactive view of technological change with the linear model of medical innovation; feedback mechanisms; explores three mechanisms by which techno-logical change may contribute to rising health-care spending; laparoscopic cholecystectomy
Gellman, 1993	Health technology assessment	J (letter)	No	No	Early involvement of Canada in HTA
George, 1995	Editor's corner: medical technology and competitiveness in the world market: reinventing the environment for innovation	J	Yes, p151	Yes, p151	Medical technology innovation; barriers to innovation; reforming the regulatory process
Goodman, 1992	It's time to rethink health care technology assessment	J	No	Yes, p343–8	HTA should be explicitly related to quality assurance, health services research, effectiveness research, medical informatics and technological innovation
Grimes, 1993	Technology follies: the uncritical acceptance of medical innovation	J	No	No	Need for on-going assessments of both new and old medical technologies
Hailey & Crowe, 1991	Health technology assessment: an Australian perspective	J	No	No	Overview of HTA in Australia
Hailey *et al*, 1990	The impact of health technology assessment	J	Yes, p225	Yes, p224–5	Measurement of the impact of HTA in terms of perceived impact on policy
Hendee, 1990	Technology transfer and cost constraints in medicine	J	No	No	Technology transfer, rising healthcare costs, cost containment, effect of cost controls and market needs
Hendee, 1991	Technology assessment in medicine: methods, status and trends	J	No	Yes, p69–73	Methods of technology assessment
Hillner *et al*, 1993	Principles of cost-effectiveness analysis for the assessment of current and new therapies	J	No	Yes, p501–6	Use of decision analysis and clinical economics to assess current and new technologies
Hoare, 1992	Tidal wave: new technology, medicine and the NHS	R	No	Yes, p2,26	Report based on the proceedings of the Caversham Conference on HTA

[*] B, book; C, conference proceedings; DP, discussion paper; J, journal article; P, paper; R, report.

Author	Title	Type*	Timing discussed?	How to assess discussed?	Focus of paper
Holohan, 1993	Assessment of new technology: regulatory agency perspective	J	No	No	Quality of evidence in assessments
Huston, 1992	Is health technology assessment medicine's rising star?	J	No	No	Rise of HTA
Hutton, 1993	Assessment of medical technology: the role of engineers	J	No	Yes, p13–14	Engineers have important inputs to HTA throughout the process
Institute of Medicine, 1985	Assessing medical technologies	B	Yes, chap 4, 7	Yes, chap 3, 7	Scope of US medical technology assessment; methods of assessment; effects of clinical evaluation on the diffusion of medical technology; reimbursement and technology assessment; opportunities for international collaboration
Jacob & Battista, 1993	Assessing technology assessment: early results of the Quebec experience	J	No	No	Independent review of the first 4 years of the Quebec Council on Health Care Technology Assessment
Jamison *et al*, 1995	Investing in health wisely: the role of needs-based technology assessment	J	No	No	Proposes essential national public health and clinical packages based on assessment of the burden of disease
Jefferson, 1996	Economic evaluations to aid decisions to conduct a trial	J (letter)	Yes, p141	Yes, p141–2	Using economic evaluation to aid making explicit society's trade-offs when funding a trial
Jennett, 1992	Health technology assessment	J	No	No	Comments on the report *Tidal Wave*
Johannesson, 1994	The concept of cost in the economic evaluation of health care: a theoretical inquiry	J	No	Yes, p675–82	The costs that should be included in an economic evaluation
Johansen, 1988	WHO concept of health technology assessment	J	No	No	WHO concept of HTA is moving towards evaluating, disseminating and monitoring results of technologies introduced to solve specific health problems
Jonsson, 1993	Economic evaluation of health care technologies	J	No	Yes, p50–4	Examines basic principles for assessing costs and benefits
Kikuchi, 1991	Status of medical engineering technology assessment in Japan	J	Yes, p12	Yes, p10–12	The development of a programme for medical engineering technology assessment in Japan
Koch, 1996	How to evaluate and implement new technologies in an era of managed care and cost containment	J	No	Yes, p799–801	How new clinical laboratory technologies can and should be chosen, evaluated and implemented
Kronborg, 1989	Randomized controlled trials in surgery	J	No	Yes, p126–7	Use of RCTs in surgery
Lantos, 1994	Ethics, randomization and technology assessment	J	Yes, p2656	Yes, p2653–6	Equipoise and RCTs
Lilford & Jackson, 1995	Equipoise and the ethics of randomization	J	No	Yes, p552–8	Discusses the ethical importance of equipoise to RCTs
Linton & Naylor, 1990	Organized medicine and the assessment of technology: lessons from Ontario	J	No	No	Attempts by organised medicine to promote the sensible introduction of new technologies
Littell, 1994	Innovation in medical technology: reading the indicators	J	No	No	Indicators of input and outputs associated with medical device innovation
Littenberg, 1992	Technology assessment in medicine	J	Yes, p425–7	Yes, p425–7	Scheme proposed for the comprehensive evaluation of medical technologies – five levels of HTA
Love, 1975	Drugs and operations: some important differences	J	No	Yes, p37–8	Argues against using routine controlled clinical trials to evaluate new operations

* *B, book; C, conference proceedings; DP, discussion paper; J, journal article; P, paper; R, report.*

Author	Title	Type*	Timing discussed?	How to assess discussed?	Focus of paper
Luce, 1995	Policy implications of modeling the cost-effectiveness of health care technologies	J	No	Yes, p1469–75	Policy implications of modelling cost-effectiveness analysis
Luce & Brown, 1995	Use of technology assessment by hospitals, health maintenance organizations, and third-party payers in the United States	J	Yes, p86	Yes, p85	The function of technology assessment in the decision-making process of healthcare providers and payers
Lumsdon, 1992	Beyond tech assessment: balancing needs, strategy	J	No	Yes, p22–6	Evaluation of healthcare technology in hospitals – considers community needs, escaping the glamour of technology, medical staff conflicts, need for formal assessment procedures, non-hospital influences on diffusion, outcomes; five technology assessment objectives outlined
MacIntyre, 1995	Efficacy assessment criteria based on risk and cost	J	No	Yes, p968–9	Three levels of efficacy assessment criteria for innovations: engineering and clinical performance assessment; physiological assessment; and clinical outcome assessment
Maisey & Lewis, 1994	NHS research and development: a research strategy for nuclear medicine	J	No	No	Work of the Diagnostics and Imaging Panel in identifying priorities for HTA for consideration by the Standing Group on Health Technology
McConnell, 1993	Six honest serving men	J	No	Yes, p33–8	Technology assessment in the context of medical-surgical nursing; questions surrounding healthcare technology pertain to the areas of need, safety, efficacy and effectiveness, economics, and social impact
McDonough, 1993	Current technology assessment programs/procedures	J	No	No	How the Office of Technology Assessment operates and some current projects
McGregor, 1994	Can our health services be saved by technology evaluation? The Quebec experience	J	Yes, p335	No	The real value of technology assessment lies in making difficult choices between different technologies; classification of technologies as experimental, innovative or accepted
Mendelson *et al*, 1995	State involvement in medical technology assessment	J	No	Yes, p93	US states technology assessment activities – Minnesota, Oregon, Washington
Menon, 1993	Technology assessment and biomedical engineering education	J	No	No	Technology assessment education and biomedical/clinical engineering
Menon & Marshall, 1996	The internationalization of health technology assessment	J	No	Yes, p46–7	HTA history, reasons for international cooperation, current international cooperative activities
Menon *et al*, 1995	Development of a health technology assessment program: the case of Alberta	J	Yes, p95	Yes, p96–7	HTA model developed by the Alberta Implementation Committee for Health Technology Assessment
Moloney & Rogers, 1979	Medical technology – a different view of the contentious debate over costs	J	No	No	To control medical technology costs requires a shift away from attempts to harness big technologies, and towards incentives to encourage more discerning use of all technologies
MRC, 1994	Health technology assessment in surgery: the role of the randomized controlled trial	R	Yes, p6–7	Yes, p2–13	Role of RCT in surgery and problems associated with it
Muller, 1991	Objective health care technology evaluation – it isn't easy	J	No	Yes, p121–4	Methods of technology assessment and how to recognise and avoid subjective biases
Murphy, 1991	Assessment process: a microscopic view	J	No	Yes, p77–82	Medical technology life cycle, phases of first stage studies, study designs

* B, book; C, conference proceedings; DP, discussion paper; J, journal article; P, paper; R, report.

Author	Title	Type[*]	Timing discussed?	How to assess discussed?	Focus of paper
Murray, 1994	Assessing genetic technologies: two ethical issues	J	Yes, p574	No	Discusses ten factors which characterise the social context of contemporary genetics; questions whether more choice is better and all improvements desirable
NHS Executive, 1996	Report of the NHS Health Technology Assessment Programme 1996	R	No	Yes, p19–20	Aims of NHS HTA programme, identifying priorities, getting work funded, progress of current projects and methods of dissemination
Nightingale, 1993	Device 'fast track' plan	J	Yes, p693	Yes, p693	US FDA plan to speed up review of critical medical devices while assuring safety and effectiveness
Ong, 1996	Lay perspective in health technology assessment	J	No	No	Nature of lay knowledge and relevance to HTA
Patail & Aranha, 1995	Role of the biomedical engineering department in William Beaumont Hospital's technology assessment process	J	No	Yes, p291–2	Assessment of medical technologies in a hospital environment; role of biomedical engineering department in the technology assessment process
Perry & Wilkinson, 1992	The technology assessment and practice guidelines forum	J	No	Yes, p289–99	A new group-judgement process, modified from the consensus development approach of the National Institutes of Health
Perry et al, 1993	Report from the Canadian Coordinating Office of Health Technology Assessment (CCOHTA)	J	Yes, p312	No	Satellite meeting to eighth ISTAHC; perspectives on technology assessment; difficulty in providing early warning of emerging technologies to policy-makers
Phelps & Parente, 1990	Priority setting in medical technology and medical practice assessment	J	No	No	Considerable differences exist in the way doctors use various medical interventions; priority setting for technology assessment
Pollock, 1989	Rise and fall of the random controlled trial in surgery	J	No	Yes, p163–8	RCT in surgery now threatened by ethical considerations
Powe & Griffiths, 1995	Clinical–economic trial: promise, problems and challenges	J	No	Yes, p377–94	Reasons for consideration of economic data collection and analysis in clinical trials; various designs and methods for gathering economic trial data
Power, 1995	Identifying health technologies that work	J	No	Yes, p205	Identifying effective practices, conducting effectiveness trials, guideline development and dissemination
Power et al, 1994	Technology assessment and public health	J	No	Yes, p566–9	Origin of HTA; its contribution to improving health decisions; techniques; putting evidence into practice; applications
Reiser, 1992	Out of chaos: a rational approach to assessing technology	J	No	Yes, p22–3	Four categories of hospital objectives for measuring technology: civic, institutional, organisational and practice goals; assessment model outlined
Rettig, 1991	Technology assessment – an update	J	No	Yes, p170–1	Reviews evolution of HTA in the 1970s and 1980s; discusses implications of creation of the Agency for Health Care Policy and Research
Rona & Beech, 1993	The process of evaluation of a new technology: genetic services and the introduction of DNA probes	J	Yes, p190	Yes, p186	Overview of process of evaluation of new technology in relation to DNA probes; barriers against application of research findings in practice
Royal, 1994	Technology assessment: scientific challenges	J	No	Yes, p505–7	Diagnostic technology assessment, levels of efficacy, marginal cost-effectiveness

[*] *B, book; C, conference proceedings; DP, discussion paper; J, journal article; P, paper; R, report.*

Author	Title	Type*	Timing discussed?	How to assess discussed?	Focus of paper
Russell, 1996	Can it work? Does it work? Research design for health technology assessment	P	No	Yes, p2–9	Methods of HTA
Russell & Grimshaw, 1995	Health technology assessment: basis of valid guidelines and test of effective implementation	B (chap)	No	Yes, p71–81	General principles of HTA, methods, explanatory and pragmatic RCTs, assessment of clinical guidelines and factors influencing validity and effectiveness
Rutten & Bonsel, 1992	High cost technology in health care: a benefit or a burden?	J	Yes, p572	Yes, p571–2	Mechanisms underlying the emergence of expensive health technology; policy implications
Rutten & Drummond, 1994	Making decisions about health technologies: a cost-effectiveness perspective	B	Yes, p27–30	Yes, chap 2–7	Role of economic appraisal in decisions about health technologies; includes three case studies
Sainfort, 1991	Evaluation of medical technologies: a generalised ROC analysis	J	No	Yes, p208–19	A generalised ROC analysis for the evaluation of medical technologies
Schulman et al, 1995	A health services approach for the evaluation of innovative pharmaceutical and biotechnology products	J	No	Yes, p1405–11	Reviews the application of health services research techniques in the assessment of potential clinical end-points
Sculpher et al, 1995	Economic evaluation in health care research: undertake it early and often	DP	Yes, p5,18	Yes, p1–18	Describes early and iterative use of economic evaluation as an integral part of research and development; the approach has four stages of economic evaluation
Shaffer & Shaffer, 1995	Support for biomedical equipment decision making	J	No	No	Hospital clinical engineering staff can provide support for biomedical equipment decision-making
Sheldon & Faulkner, 1996	Vetting new technologies: those whose efficacy and safety have not been established will now be registered and evaluated	J	No	No	Questions raised by the establishment of SERNIP
Sheldon et al, 1992	Health technology assessment	J	No	No	Information on bulletin, *Effective Health Care*
Sibbald et al, 1993	New technologies, critical care, and economic realities	J	No	No	HTA in critical care services
Simpson, 1993	Managing innovative technology	J	No	No	Guidelines for managing innovative technologies
Sisk & Glied, 1994	Innovation under federal health care reform	J	Yes, p93–4	No	In a climate of cost containment, systematic evaluation of new technology is vital to identify and expand coverage to worthwhile innovations
Sloan et al, 1986	Diffusion of surgical technology: an exploratory study	J	No	No	Empirical analysis of the diffusion patterns of five surgical procedures
Smith, 1994	Towards a knowledge based health service: priorities are set for health technology assessment	J	No	No	Describes the NHS HTA programme
Smith et al, 1994	Role of economic appraisal in health technology assessment: the Australian case	J	Yes, p1654, 1661–2	Yes, p1653–62	Examines role and importance of economic appraisal of health technology in Australia – includes eight case studies
Sniderman, 1996	Governance of clinical trials	J	No	Yes, p1387–8	Argues that local hospital ethical committees should play a greater part in the supervision of clinical trials
Solomon & McLeod, 1993	Clinical assessment of biomedical technology	J	No	Yes, p301–6	Accuracy, reliability and validity in assessment of diagnostic technologies

* B, book; C, conference proceedings; DP, discussion paper; J, journal article; P, paper; R, report.

Author	Title	Type*	Timing discussed?	How to assess discussed?	Focus of paper
Standing Group on Health Technology, 1994	1994 report	R	No	Yes, p8	First report of SGHT
Steering Committee on Future Health Scenarios, 1987	Anticipating and assessing health care technology (vol 1)	B	Yes, chap 3	Yes, chap 3, 7	Includes: development and diffusion of healthcare technology; identifying and assessing healthcare technology; anticipated changes in healthcare technology
Stephenson, 1995	Medical technology watchdog plays unique role in quality assessment	J	No	No	Outlines role of ECRI in HTA
Stevens *et al*, 1995	'Quick and clean': authoritative health technology assessment for local health care contracting	J	Yes, p37–8	Yes, p37–42	Seven-stage model of evaluation
Stiefel & Rizkalla, 1995	The elements of a complete product evaluation	J	No	Yes, p482–8	Outlines a formal methodology for completing a product evaluation
Stocking, 1988	Factors influencing the effectiveness of mechanisms to control medical technology	B (chap)	Yes, p20–1, 24–5	No	Problems in technology assessment; mechanisms for control of uptake; groups who influence the success of the mechanisms
Sussman, 1991	Financial considerations in technology assessment	J	No	No	Covers: growth in healthcare technology; resource planning and technology acquisition; entering the high-tech market; financing long-term acquisition strategies
Szczepura, 1993	Health care technology assessment in Europe: training for the future	B	Yes, chap 2	Yes, chap 4	Results of European study to obtain views of healthcare providers, funders, manufacturers, professional associations, academic institutes and policy-making bodies on their need for expertise and competence in HTA
Szczepura & Cooke, 1993	Softly, softly	J	No	No	Describes changes to enable NHS managers to harness HTA findings more effectively
Szczepura & Kankaanpaa, 1996	Assessment of health care technologies: case studies, key concepts and strategic issues	B	Yes, p33,227	Yes, chap 3–13	Discusses key concepts and strategic issues in HTA – includes eight case studies
Task Force on Principles for Economic Analysis of Health Care Technology, 1995	Economic analysis of health care technology: a report on principles	J	No	Yes, p61–8	Guidelines for researcher independence and guidelines for reporting economics outcomes research
Teplensky, 1995	Hospital adoption of medical technology: an empirical test of alternative models	J	No	No	Reviews three common explanations for medical technology adoption – profit maximisation, technological pre-eminence, and clinical excellence – and incorporates them into a composite model
Thacker & Berkelman, 1986	Surveillance of medical technologies	J	No	Yes, p369–71	Argues for surveillance of technologies as they diffuse into practice
Thibault, 1992	Evaluating medical technology in the 1990s	J	No	Yes, p266	Evaluation of practice patterns in a medical intensive care unit
Thornbury & Fryback, 1992	Technology assessment – an American view	J	No	Yes, p147–54	Proposes a hierarchical model to enhance understanding of the interrelations of different aspects of technology assessment

* B, book; C, conference proceedings; DP, discussion paper; J, journal article; P, paper; R, report.

Author	Title	Type*	Timing discussed?	How to assess discussed?	Focus of paper
Tugwell, 1995	Technology assessment: old, new, and needs-based	J	No	Yes, p656–60	Needs-based technology assessment focuses on the population-wide impact of an intervention
Tuman & Ivankovich, 1993	High-cost, high-tech medicine: are we getting our money's worth?	J	No	Yes, p168–76	Reviews principles of economic and cost–benefit analyses, outcome benefit studies, and decision-threshold analysis
US Congress: Office of Technology Assessment, 1994	Identifying health technologies that work: searching for evidence	B	No	Yes, chap 3, 5	Includes: tools for effectiveness research; issues in improving effectiveness research; cost-effectiveness analysis; the federal role in HTA
van Gennip & Gremy, 1993	Challenges and opportunities for technology assessment in medical informatics	J	Yes, p182–3	No	Report of a MEDINFO '92 workshop on challenges and opportunities for technology assessment in medical informatics
Werko & Banta, 1995	Report from the EUR-ASSESS Project	J	No	No	How EUR-ASSESS is trying to develop a coordinated approach to health technology assessment in Europe
Wheeler *et al*, 1985	Technology: a strategic factor in hospital planning	J	No	No	Hospital administrators must be able to *et al,* assess technology and be willing to adapt to rapid change
Whitmore, 1993	Who, what, where, when, why and how: technology assessment in a hospital setting	J	No	No	Role of biomedical engineering in technology assessment
Whittenburg, 1995	A program proposal for new technology assessment	J	No	Yes, p391–9	Describes a programme for technology assessment, its evaluation and implementation
Williams, 1993	Technology assessment in clinical laboratory science: becoming involved in a critical process	J	No	Yes, p202–5	Methods of HTA
Wilson, 1995	Case study: practical tools for improving needs-based health management and technology assessment	J	No	Yes, p709–15	Primary Health Care Management Advancement Program; how it enables health management teams to assess health services
Winkler *et al*, 1986	Popular press coverage of eight National Institutes of Health consensus development topics	J	No	No	Analysis of popular press coverage of eight National Institutes of Health consensus development conference topics

* B, book; C, conference proceedings; DP, discussion paper; J, journal article; P, paper; R, report.

Appendix 5

Research team experience

Jane Bower worked for 15 years as a molecular geneticist, at Edinburgh and Stanford Universities, and most recently at the MRC Human Genetics Unit where she led a project on the isolation of genes associated with inherited retinal disorders. She then moved to researching the process of management of new biomedical technology development, supported as principal investigator by the Royal Society, the Japan Society for the Promotion of Science, the Leverhulme Trust, the Nuffield Foundation (in a study of the development of genetic manipulation technology in the NHS) and most recently by the Economics and Social Research Council in a study of biomedical innovation in the NHS (with Lorna McKee). She is currently a senior lecturer in the Department of Management Studies, Aberdeen University.

Lorna McKee has a DPhil in Sociology (University of York) and has extensive research experience in healthcare management, management of change and primary care. She has worked as principal investigator on studies funded by the Economics and Social Research Council, Equal Opportunities Commission, Health Education Authority, and Scottish Office Home & Health Department. She is also employed as a research consultant to several NHS Health Authorities and Boards. She is currently Director of Research in the Department of Management Studies, University of Aberdeen, with current projects on the roles of Clinical Directors in Scottish Trusts (Economics and Social Research Council) and primary care innovation (Scottish Office Home & Health Department).

Adrian Grant is Director of the Health Services Research Unit, a national centre for Scotland

core funded by the Chief Scientist's Office of the Scottish Office Department of Health. Within its programme on HTA, the unit is working on MIS and telemedicine applications. Before coming to Aberdeen, he directed a programme of HTA in the perinatal field, coordinating several large-scale international trials, including the MRC multicentre trial of CVS. He also has experience of systematic reviews within the NHS research and development programme and the Cochrane Collaboration.

John Cairns is the Director of the Health Economics Research Unit, another national unit for Scotland. He was a member of the SCOTMEG/CRAG working party on High Technology Equipment, a member of the Steering Group for MRC Genetic Approaches to Human Health, and is on a Scottish Office Department of Health Cancer Genetics subcommittee. The Health Economics Research Unit is currently undertaking MRC-funded research on health economic aspects of MIS.

John Brebner has been involved with the development of telemedicine for 15 years, as a practitioner and a researcher. His PhD research focused on this topic and he has an extensive international network of contacts who are currently involved in developing applications of this technology.

Graham Mowatt, the Research Fellow engaged on the project, is a chartered librarian, and has extensive experience of undertaking manual and automated literature searches, particularly in biomedical projects. He possesses an MBA degree and was previously employed by Grampian Health Board for 4 years on a health service project on technological innovation in primary care.

Copies of this report can be obtained from: