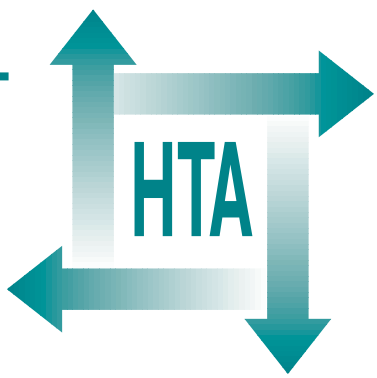
Health Technology Assessment 1999; Vol. 3: No. 18

Review

A systematic literature review of spiral and electron beam computed tomography: with particular reference to clinical applications in hepatic lesions, pulmonary embolus and coronary artery disease

E Berry
S Kelly
J Hutton
KM Harris
P Roderick
JC Boyce
J Cullingworth
L Gathercole
PJ O'Connor
MA Smith

Health Technology Assessment
NHS R&D HTA Programme



Standing Group on Health Technology

Current members

Chair:
Professor Sir Miles Irving,
Professor of Surgery, University of Manchester, Hope Hospital, Salford

Professor Martin Buxton,
Professor of Economics, Brunel University

Professor Francis Creed,
School of Psychiatry & Behavioural Sciences, University of Manchester

Professor Charles Florey,
Department of Epidemiology & Public Health, Ninewells Hospital & Medical School, University of Dundee

Professor John Gabbay,
Director, Wessex Institute for Health Research & Development

Professor Sir John Grimley Evans,
Department of Geriatric Medicine, Radcliffe Infirmary, Oxford

Dr Tony Hope,
The Medical School, University of Oxford

Professor Richard Lilford,
Regional Director, R&D, West Midlands

Dr Jeremy Metters,
Deputy Chief Medical Officer, Department of Health

Professor Maggie Pearson,
Regional Director of R&D, NHS Executive North West

Mr Hugh Ross,
Chief Executive, The United Bristol Healthcare NHS Trust

Professor Trevor Sheldon,
Director, NHS Centre for Reviews & Dissemination, University of York

Professor Mike Smith,
Dean of Research, Faculty of Medicine, Dentistry, Psychology and Health, University of Leeds

Dr John Tripp,
Department of Child Health, Royal Devon & Exeter Healthcare NHS Trust

Professor Tom Walley,
Department of Pharmacological Therapeutics, University of Liverpool

Dr Julie Woodin,
Chief Executive, Nottingham Health Authority

Professor Kent Woods
(**Chair Designate**),
Regional Director of R&D, NHS Executive, Trent

Past members

Dr Sheila Adam,
Department of Health

Professor Angela Coulter,
Director, King's Fund, London

Professor Anthony Culyer,
Deputy Vice-Chancellor, University of York

Dr Peter Doyle,
Executive Director, Zeneca Ltd, ACOST Committee on Medical Research & Health

Professor John Farndon,
Professor of Surgery, University of Bristol

Professor Howard Glennester,
Professor of Social Science & Administration, London School of Economics & Political Science

Mr John H James,
Chief Executive, Kensington, Chelsea & Westminster Health Authority

Professor Michael Maisey,
Professor of Radiological Sciences, Guy's, King's & St Thomas's School of Medicine & Dentistry, London

Mrs Gloria Oates,
Chief Executive, Oldham NHS Trust

Dr George Poste,
Chief Science & Technology Officer, SmithKline Beecham

Professor Michael Rawlins,
Wolfson Unit of Clinical Pharmacology, University of Newcastle-upon-Tyne

Professor Martin Roland,
Professor of General Practice, University of Manchester

Professor Ian Russell,
Department of Health Sciences & Clinical Evaluation, University of York

Dr Charles Swan,
Consultant Gastroenterologist, North Staffordshire Royal Infirmary

Details of the membership of the HTA panels, the NCCHTA Advisory Group and the HTA Commissioning Board are given at the end of this report.

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.hta.ac.uk). A fully searchable CD-ROM is also available (see below).

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

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

You can order HTA monographs from our Despatch Agents:

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

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

Contact details are as follows:

HTA Despatch
c/o Direct Mail Works Ltd
4 Oakwood Business Centre
Downley, HAVANT PO9 2NP, UK

Email: orders@hta.ac.uk
Tel: 02392 492 000
Fax: 02392 478 555
Fax from outside the UK: +44 2392 478 555

NHS libraries can subscribe free of charge. Public libraries can subscribe at a very reduced cost of £100 for each volume (normally comprising 30–40 titles). The commercial subscription rate is £300 per volume. Please see our website for details. Subscriptions can only be purchased for the current or forthcoming volume.

Payment methods

Paying by cheque
If you pay by cheque, the cheque must be in **pounds sterling**, made payable to *Direct Mail Works Ltd* and drawn on a bank with a UK address.

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

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

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

Please use the form on the HTA website (www.hta.ac.uk/htacd.htm). Or contact Direct Mail Works (see contact details above) by email, post, fax or phone. *HTA on CD* is currently free of charge worldwide.

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

A systematic literature review of spiral and electron beam computed tomography: with particular reference to clinical applications in hepatic lesions, pulmonary embolus and coronary artery disease

E Berry[1*]
S Kelly[1]
J Hutton[2]
KM Harris[3]
P Roderick[4]
JC Boyce[1]
J Cullingworth[3]
L Gathercole[3]
PJ O'Connor[3]
MA Smith[1]

[1] Medical Physics and Engineering, and Centre of Medical Imaging Research, University of Leeds/Leeds Teaching Hospitals NHS Trust, Leeds General Infirmary, Leeds, UK
[2] MEDTAP International Inc., London, UK
[3] Department of Radiology, Leeds Teaching Hospitals NHS Trust, Leeds General Infirmary, Leeds, UK
[4] Wessex Institute of Health Research and Development, Southampton University, Southampton, UK

[*] Corresponding author

Competing interests: none declared

Published October 1999

This report should be referenced as follows:

Berry E, Kelly S, Hutton J, Harris KM, Roderick P, Boyce JC, *et al.* A systematic literature review of spiral and electron beam computed tomography: with particular reference to clinical applications in hepatic lesions, pulmonary embolus and coronary artery disease. *Health Technol Assess* 1999;**3**(18).

Health Technology Assessment is indexed in *Index Medicus*/MEDLINE and *Excerpta Medica*/EMBASE. Copies of the Executive Summaries are available from the NCCHTA web site (see overleaf).

NHS R&D HTA Programme

The 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 Diagnostics and Imaging Panel and funded as project number 94/28/01.

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. The editors wish to emphasise that funding and publication of this research by the NHS should not be taken as implicit support for the recommendations for policy contained herein. In particular, policy options in the area of screening will be considered by the National Screening Committee. This Committee, chaired by the Chief Medical Officer, will take into account the views expressed here, further available evidence and other relevant considerations.

Reviews in *Health Technology Assessment* are termed 'systematic' when the account of the search, appraisal and synthesis methods (to minimise biases and random errors) would, in theory, permit the replication of the review by others.

Series Editors: Andrew Stevens, Ruairidh Milne and Ken Stein
Monograph Editorial Manager: Melanie Corris

The editors have tried to ensure the accuracy of this report but cannot accept responsibility for any errors or omissions. They would like to thank the referees for their constructive comments on the draft document.

ISSN 1366-5278

Published by Core Research, Alton, on behalf of the NCCHTA.
Printed on acid-free paper in the UK by The Basingstoke Press, Basingstoke.

Copies of this report can be obtained from:

The National Coordinating Centre for Health Technology Assessment,
Mailpoint 728, Boldrewood,
University of Southampton,
Southampton, SO16 7PX, UK.
Fax: +44 (0) 23 8059 5639 Email: hta@soton.ac.uk
http://www.hta.nhsweb.nhs.uk

Contents

List of abbreviations

A	intercept of SROC (see chapter 3)
ANOVA	analysis of variance†
B	gradient of SROC (see chapter 3)
BIDS	Bath Information and Data Services
CAC	coronary artery calcium
CAD	coronary artery disease
CCT	conventional computed tomography†
CT	computed tomography/tomographic
CTA	computed tomography angiography
CTAP	computed tomography arterial portography
CTDI	computed tomography dose index
D	vertical axis of SROC
DP	delayed phase
EBCT	electron beam computed tomography
EWLS	equally weighted least squares (see chapter 3)
FDA	Federal Drug Administration†
FN	false-negative†
FP	false-positive†
FPR	false-positive rate = 1 – specificity
FPR2	statistic resulting from the addition of 0.5 to each of TP, FN, FP and TN
HAP	hepatic arterial phase
HCC	hepatocellular carcinoma
ISI	Institute of Scientific Information
LiB	lithium borate†
LiF	lithium fluoride†
MeSH	medical subject heading
MRI	magnetic resonance imaging
MSAD	mean scan average dose
N/App	not applicable†
N/P	not performed†
NPV	negative predictive value
N/S	not stated†
N/Sig	not statistically significant†
OCLC	Online Computer Library Center
OR	odds ratio
OR2	statistic resulting from the addition of 0.5 to each of TP, FN, FP and TN
PE	pulmonary embolism/embolus/emboli
PPV	positive predictive value
PTCA	percutaneous transluminal coronary angioplasty
PVP	portal venous phase
Q*	Q star (see chapter 3)
QoL	quality of life
QALY	quality-adjusted life year
RARE	rapid acquisition with relaxation enhancement
ROC	receiver operator characteristic
RR	robust resistant (see chapter 3)
S	horizontal axis of SROC
SCT	spiral computed tomography
Se	standard error
SROC	summary receiver operator characteristic
TLD	thermoluminescent dosimeter†
TN	true-negative†
TP	true-positive†
TPR	true-positive rate = sensitivity
TPR2	statistic resulting from the addition of 0.5 to each of TP, FN, FP and TN
Un	unenhanced†
US	ultrasound
VP	ventilation/perfusion

† Used only in tables and figures

Executive summary

Objectives

The aim of this review was to identify publications relating to the use of spiral and electron beam computed tomography (CT), in order to draw conclusions about the effectiveness of latest generation CT devices. The Fineberg evaluative framework was used, with publications sought for all clinical applications at the levels of health economics, patient outcome and therapeutic impact. For diagnostic impact and diagnostic performance, specific clinical uses were selected: the investigation of liver lesions using spiral CT (SCT); the investigation of pulmonary embolism (PE) using SCT; and the diagnosis and prediction of coronary artery disease (CAD) using electron beam CT (EBCT).

Methods

Data sources

MEDLINE and BIDS-ISI formed the basis of the literature search. Other electronic resources searched included the Cochrane Library, EMBASE, Inside Information Plus, the System for Information on Grey Literature, and FirstSearch. Bibliographic listings of all retrieved articles were handsearched. In addition, manufacturers were contacted with a request for unpublished information.

Study selection

Study selection was a three-stage process using predefined inclusion and exclusion criteria. Non-English language papers were excluded. In the assessment of health economics, patient outcome, therapeutic impact and diagnostic impact, important validity criteria were identified and study compliance noted, but studies were not excluded on this basis. For studies of diagnostic performance, a checklist approach was used to record the risk of bias and methodological differences. Quantitative synthesis was performed only in the case of EBCT for CAD. The results of the checklists were incorporated into the data synthesis to assess the influence of biases and factors on the reported diagnostic performance.

Data extraction

Data extraction forms were used. Numerical values for the completion of 2×2 contingency tables were extracted when possible. Descriptive summaries were prepared for the other types of study when quantitative analysis was not feasible.

Data synthesis

Qualitative synthesis was used for the studies of health economics, patient outcome, therapeutic impact and diagnostic impact.

The results of studies of diagnostic performance for one out of three clinical applications were synthesised into summary receiver operator characteristic (SROC) curves. Study validity was investigated by using regression techniques to incorporate biases and factors into the quantitative analysis. For each bias or factor found to be significant ($p < 0.05$) its influence on the diagnostic performance was illustrated with SROC curves.

Results

- There was very little health economics evidence relating to either SCT or EBCT. Only one study approached acceptable standards for economic evaluation.
- Patient outcome or therapeutic impact was adequately addressed by nine studies: three for SCT and six for EBCT.
  - The indications from three studies on EBCT that addressed patient outcome are that EBCT (as opposed to no EBCT) may improve outcome in a variety of clinical scenarios, but results are by no means conclusive.
  - Insufficient information was found for qualitative synthesis regarding the therapeutic impact of the use of SCT and EBCT. It is likely that results will depend upon the clinical application and the comparator investigation.
- No studies were identified that were designed specifically to address the diagnostic impact of either modality.
  - The included studies that compared the detection performance of SCT for liver lesions with conventional CT found that SCT performed better, but disagreed about the size of lesion that was best detected.
  - There were conflicting findings when comparing SCT for the detection of liver

lesions with alternative modalities other than conventional CT.
  - Insufficient information was found for qualitative synthesis that compared results from other modalities with SCT in PE, and with EBCT in CAD.
- From 1515 articles that satisfied the preliminary inclusion criteria on diagnostic performance in the three clinical application areas, 49 satisfied the inclusion criteria for the qualitative review, and 7 for quantitative analysis.
  - Four articles were included that measured the diagnostic performance of SCT applied to liver lesions against a gold standard, but no conclusions could be drawn from them because there was great variation among the individual studies. Those that compared performance with conventional CT showed an increase in the number of lesions detected by SCT of the order of 10%.
  - SCT detection performance for PE is better for the central vessels alone than for both central and peripheral vessels together.
  - Insufficient information for qualitative synthesis was found regarding the prediction of asymptomatic CAD by using EBCT.
  - Six studies on the diagnosis of symptomatic CAD using EBCT had a low specificity (high false-positive rate). The most likely role for EBCT is in excluding obstructive CAD in the older population.
  - Interobserver and intraobserver reproducibility of EBCT for CAD is acceptable, but interexamination reproducibility is not.
- A total of 11 studies of radiation dose were included in the review.
  - Of four studies comparing SCT dose with that of conventional CT, the general agreement was of an insignificant increase in dose over conventional CT, with SCT offering the potential for reducing the dose by increasing pitch.
  - Insufficient information for qualitative synthesis was found regarding the radiation dose in EBCT.

Conclusions

- MEDLINE and BIDS-ISI are comprehensive sources of references in this subject area.
- There is no strong evidence about any aspect of the use of latest generation CT at the health economics, patient outcome or therapeutic impact levels.
- SCT detects liver lesions that are not seen with conventional CT.
- EBCT has a low specificity when applied to the diagnosis of symptomatic CAD.
- While the evidence suggests EBCT use for population studies, it does not support its use to track CAD progression in individuals.
- The introduction of SCT will not cause a significant increase in radiation dose compared with similar examinations performed with conventional CT.

Recommendations for research

Methodological:

- methodological research into the effect of searching only the major electronic databases and into factors that make publication bias less likely
- continued collaboration between reviewers in fields that are lacking in randomised controlled trials regarding the assessment of study quality
- further research into SROC methodology when applied to tests requiring unequal sensitivity and specificity
- horizon scanning to identify developments such as SCT that may rapidly become accepted before assessment has been performed
- the encouragement of imaging scientists both to perform better designed studies and to ensure that descriptions published in the literature are comprehensive.

Topic related:

- updating of this review, especially with regard to long-term follow-up of EBCT in asymptomatic individuals, and for SCT in PE
- a multicentre study of SCT for liver lesions, using a group of affected patients to investigate optimum automatic protocols, and with careful control of intrinsic factors
- studies on the reduction of contrast medium dose, using automatic injection protocols, concentrating on the detectability of liver lesions rather than maximal parenchymal enhancement
- a systematic literature review on the clinical relevance of subsegmental PE
- after the review of subsegmental PE, research using decision analytical modelling to compare a variety of diagnostic strategies including ventilation/perfusion, SCT, magnetic resonance imaging and pulmonary angiography
- new studies designed specifically to measure diagnostic impact, therapeutic impact and patient outcome
- the use of decision-modelling techniques to combine outcome and cost data from a variety of sources and new studies.

Chapter 1
Background

Although this systematic review deals with a specific area of medical imaging, it is an example of a general situation that exists in the area of diagnostics and imaging that poses particular problems for the assessment of health technology. Scientific breakthroughs can produce new 'devices' that can have a significant clinical impact. Such devices often continue to improve over the years as a result of related technological advances. Their performance continues to change and we can consider them as 'evolving technologies'. The methodological problems associated with the assessment of these technologies have been the subject of a research programme within the NHS R&D initiative.[1] Although this research has implications for the future, the existing situation is such that there is a problem associated with the assessment of evolving health technology because there is incomplete coverage of health technology assessment topics in current research activity and in published work. Computed tomography (CT) is a particular example of this situation. The basic technique has been in clinical use since the early 1970s, when it had a major impact and its clinical value was widely accepted. Subsequently, the technology has evolved, partly, it must be said, owing to commercial pressures to maintain competitive advantage. Thus there exists within the field of medical imaging the understanding that the value of CT was assessed many years ago, with the result that there is not a great deal of concern related to its cost-effectiveness or impact on patient outcome.

The evolution of medical devices occurs in parallel with general technological developments that become incorporated into the devices. A consequence of general developments in technology (computers are a good example) is that the constituent components of a device could reduce in cost. However, when medical devices are considered, it is usually the case that general technological developments are not translated into significant cost reductions. There is a tendency to increase the performance and sophistication of devices such that their cost is maintained, or indeed increased, in the long term. A key question for health technology assessment is to determine the appropriate level of performance that is required by such devices for clinical effectiveness, and the resulting economic consequences.

The equipment

X-ray CT is an imaging technique that was developed in the early 1970s and eagerly utilised for its ability to provide cross-sectional images of the body. The technology of CT has continued to develop, resulting in reduced scanning times for a single slice. The developments are sometimes referred to as first-, second-, third- (and so on) generation CT systems. Although each subsequent generation was associated with a technical advance, the most recent technical advances have probably been the greatest: the development of helical CT[2] and electron beam (or ultrafast) CT (EBCT).[3]

Helical CT is commonly referred to as spiral CT (SCT) and, although the term is geometrically incorrect, we will use it because it is widely understood. We will refer to SCT and EBCT under the general term 'latest generation CT'. Although there are many publications that describe the technical performance of imaging modalities, they are not included in this review, where we have considered the overall clinical effectiveness and the performance of specific clinical applications of this technology.

SCT

SCT images are acquired by continuous scanning as the patient moves through the scanner. Conventional (or incremental) machines acquire slices in series, with no relative longitudinal motion between the table holding the patient and the imaging equipment during data acquisition. SCT was made possible as a consequence of technical developments in both slip ring and X-ray tube technology. This new mechanism allows very fast acquisition of the contiguous slices used for the reconstruction of three-dimensional anatomy; a typical scan time is between 20 and 60 seconds. If the subject suspends respiration for this period, the image quality is improved because of the lack of artefacts from respiratory motion. There is the possibility of reconstructing slices at arbitrary intervals with the aim of better centring on focal lesions. The image quality obtained is also a function of the acquisition parameters chosen.[4] SCT scanners are now available from all the major medical imaging manufacturers.

SCT is now replacing conventional CT for many applications, simply because it is so much faster and allows volumetric studies to be obtained in the space of a single breath hold. This modality has found favour for CT angiography (CTA) because its speed means that reduced volumes of injected contrast material may be used and more images acquired during the period of peak enhancement.[5]

EBCT

EBCT has even shorter acquisition times than SCT, of the order of 50–100 ms. This has been achieved by using an electron beam to generate the X-rays; there is no moving gantry because the X-rays are directed by scanning the electron beam. EBCT systems are considerably less common than SCT. The majority of existing systems are manufactured by a single company, Imatron Inc. Imatron have registered the term 'Ultrafast CT' as a trade name for their company's product. The term 'electron beam CT' is more generic and further distinguishes the modality from the increasingly fast acquisitions available from scanners with moving gantries. There are 99 EBCT scanners world-wide; one is in the UK.

Effectiveness of latest generation CT devices

It has been mentioned above that earlier generations of CT devices contained technological advances that resulted in a faster scan time. SCT and EBCT are faster than the earlier generations of CT scanners. As a result, there could be issues of cost-effectiveness associated with the potentially increased throughput owing to the reduction in scanning time. However, perhaps of more importance is the level to which the scan time has been reduced. In SCT, the scan time has now been reduced so that it is possible to image a complete organ during a breath hold and, if necessary, to repeat this soon afterwards. This offers potential opportunities for diagnostic effectiveness that were not present in earlier generation CT scanners. Similarly, the short acquisition time of EBCT is such that imaging within the heart can be undertaken, which, again, was not possible with the earlier generation scanners. Thus, when considering the effectiveness of latest generation CT scanners, we must consider two distinct categories: (1) increased effectiveness in conventional applications; and (2) new applications for which there may be evidence of clinical effectiveness and cost-effectiveness.

When considering the effectiveness of diagnostic devices, it is valuable to use the hierarchy that was initially proposed by Fineberg[6] and others.[7,8] Different healthcare professionals and consumers would look for evidence of effectiveness at different levels of the hierarchy.

The classification outlined in *Figure 1* has been used for this review of latest generation CT scanners. When this hierarchy is employed, there is a general tendency to commence a review at the level of diagnostic performance and then progress through the different levels, ending with the economic impacts. This is entirely appropriate when a systematic review is being undertaken for an imaging device in association with a specific clinical question. In the situation where the device in general is being reviewed, as is the case with latest generation CT scanners, it is more appropriate to review the subject in the reverse order, starting with the economic impacts and ending with diagnostic performance. This is applicable when there could be general issues of effectiveness affecting a range of applications and where the evidence of effectiveness in specific areas may be of secondary importance.

The shape of the hierarchy as presented in *Figure 1* is deliberate. If we were to consider the quantity of research publications that were present at each level, it would appear very much pyramidal in shape, with the largest number concerned with diagnostic performance and only a small number with health economics. Such variation in the numbers of publications is, in our experience, common within the field of medical imaging.

In the following paragraphs, the types of study that might be performed at the different levels of the hierarchy are discussed.

Health economics

The focus of economic evaluation is on resource use and benefits to patients that may be realised in a routine healthcare delivery situation. In other words, the external validity of studies is more important than the internal validity, which means that randomised controlled trials have drawbacks as vehicles for economic evaluation. Nevertheless, the criteria for judging economic studies can still be grouped into four main categories: study design; data collection; analysis; and interpretation of results.

Patient outcome

A change in patient outcome may result from the combination of the next two levels in the pyramid: therapeutic and diagnostic impact. The follow-up period required to verify this will vary with the

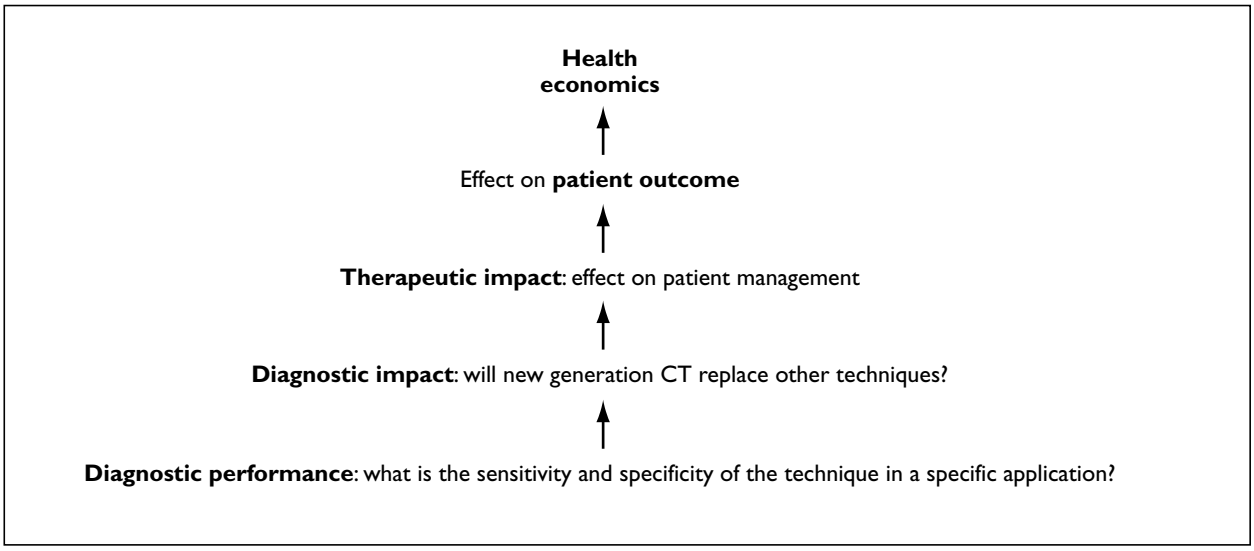


FIGURE 1 *Fineberg hierarchical classification scheme for studies of diagnostic or staging performance*

disease site. The outcome measure chosen must be appropriate to the question being addressed and the analytical approach used. If the risk of morbidity and mortality associated with the tests under study differ, then the outcome measure must be able to incorporate these effects. The timing and frequency of outcome assessment must also take this into account.

Many studies do not record systematically the impact of a new staging or diagnostic approach on actual diagnoses or on subsequent therapeutic decisions. This makes very difficult the assessment of the potential impact on outcome, regardless of the outcome measure chosen. Many clinical studies of therapeutic intervention stop short of measuring actual outcomes. For example, the success of cancer therapy is often judged on the tumour response rate, without long-term follow-up to see whether patients with a good response to treatment actually survive longer. Survival is not the only aspect of outcome of interest to patients. A large amount of published literature has built up on the assessment of quality of life (QoL) during and after treatment.[9]

Some researchers have developed guidelines for judging the relevance and validity of QoL studies.[10] These are helpful but relate only to one aspect of outcome measurement. There is a hierarchy of patient outcomes, as described by Fries *et al.*[11] and discussed in some of the economics guidelines (e.g. by Drummond and Jefferson[12]). Economists working on the evaluation of health have developed measures of outcome that combine both the change in the quantity (survival) and the quality of life, the most frequently used being the quality-adjusted life year (QALY).[13] There is still much disagreement and debate about the best way to assess the impact on QoL for such outcome measures. For example, should instruments be disease specific or generic? Should respondents be patients with the condition or representative samples of the general population? The following could be regarded as a widely accepted ranking of outcome measures, progressing from the least useful to the most broadly applicable:

- intermediate clinical outcomes (e.g. tumour response rates or number of true-positive diagnoses)
- final clinical outcomes (survival rates)
- cumulative clinical outcomes (life years saved)
- patient assessed outcomes (QoL)
  - patient satisfaction
  - disease-specific QoL scales
  - generic QoL scales
- patient preference measures (combining QoL and survival)
  - QALYs
  - healthy-year equivalents
- monetary values of patient benefits
  - willingness to pay.

The final category of monetary measures would allow economists to carry out a full cost–benefit analysis of health care with costs and benefits in monetary terms. In healthcare systems where patients do not buy care directly, willingness to pay has to be elicited by indirect means such as conjoint analysis. These approaches have been widely used in other areas of economics, such as transport and the environment, and are increasingly the subject of new research in health economics.

Their application in the imaging field has been limited to assessments of the acceptability of different tests to patients (e.g. willingness to pay for more expensive tests that involve less risk or discomfort), but their wider potential has been recognised.[14]

Therapeutic impact

This is defined in terms of changes in the clinical management of patients as a result of diagnosis by a different modality. This can involve changes between curative and palliative therapy or surgical and medical management, or a faster introduction of the same therapy. As for diagnostic impact, the basic factors of good study design are important, including initial randomisation. Sufficiently extended follow-up to allow the observation of changes in management is preferable. Studies often record intentions to treat only, particularly if patients have received two tests and clinicians are asked to assess the impact of each one independently.

Diagnostic impact

The focus of the analysis at this stage is whether the use of the new diagnostic technology leads to any patient receiving a different diagnosis. Diagnostic impact can also be defined in terms of the confidence of clinicians in their diagnoses. A more confident diagnosis can have two effects: active therapy may be undertaken more quickly, and fewer confirmatory, duplicate diagnostic tests may be used, thus reducing the cost of the diagnostic process.

To assess diagnostic impact, studies must assess how the results of tests are used by clinicians in reaching a diagnosis, and how they fit into a sequence of clinical decisions. This requires a more pragmatic or naturalistic design than that of an experimental study addressing diagnostic performance. Studies of diagnostic impact can be designed so that a study group receives the test under investigation and a control group does not. Randomisation is still desirable to prevent selection bias, but blinding the clinician to the source of information is usually not possible, nor is it necessarily desirable. In situations where tests are complementary rather than directly substitutive, the problem may be to determine the optimal sequence of testing. This can be done by randomising patients between two predetermined access routes to the first test and allowing clinicians to request the second test if desired. The latter may better reflect how the tests may subsequently be used in routine practice but does introduce the possibility of selection bias for the second test.

In the absence of studies designed specifically to evaluate diagnostic impact, those that use comparative technologies in a diagnostic performance study design can be used as a secondary standard. This will provide an experimental comparison between the performance statistics of the competing tests, but it does not supply any information regarding the subsequent impact of replacing the existing technology. The possibility of the tests being complementary may be overlooked.

Diagnostic performance

To assess diagnostic accuracy, the comparison should be made against a gold standard reference test, which should be applied to all subjects. Study designs at this level need to be rigorous and free from bias to ensure validity. Optimally, to demonstrate diagnostic performance, subjects should be allocated randomly to the study, with blinding between test and reference. However, in medical imaging, many studies do not, or cannot, adhere to this design. Imaging studies often address the clinical effectiveness of the technology and, in this scenario where the performance in clinical routine is being evaluated, the appropriate study design is not so well defined. There is much literature on study design; an article by the authors of this review has been published, which describes the potential biases that are prevalent in performance studies of medical imaging modalities.[15]

Technical performance

One technical aspect – radiation dose – will be covered in this review, because it links directly to the performance of the technology in clinical routine. In the commissioning brief for this review it was pointed out that conventional CT scans are responsible for 20% of medical X-ray exposure in the general population, while comprising only 2% of diagnostic examinations. These figures will be undergoing change as patterns of use change, but any future analysis into the costs and benefits of these technologies will need to include the likely impact of radiation dose.

Specific clinical applications

As the focus of this review was device related rather than application specific, all literature related to (1) health economics, (2) patient outcome and (3) therapeutic impact was reviewed for all diagnostic applications. For diagnostic impact and, in particular, for diagnostic performance, it was necessary to select specific clinical applications.

To investigate diagnostic performance or accuracy, we have focused on studies comparing the performance of latest generation CT with a gold standard. In order rigorously to compare performance with other modalities, such as magnetic resonance imaging (MRI) or incremental CT, it would be necessary to perform a series of separate systematic reviews to determine the absolute performance of each modality. This was clearly beyond our remit for this review. Instead, we have searched for studies that perform a self-contained comparison, by comparing both latest generation CT and an alternative modality to the same gold standard on the same group of patients.

Selection of diagnostic applications

For SCT, the choice of clinical applications was made by a survey of experts in the UK, with reference to the research priorities of the Health Technology Assessment Programme and considering other developing imaging modalities that might be reviewed separately in the application area. Full details of the methodology are given chapter 3. Briefly, workers at 19 UK centres with access to SCT were asked to indicate the three diagnostic applications that they considered were the most important. The survey results are given in *Figure 2*. The list of centres was supplied by the British Institute of Radiology.

The diagnostic applications that were selected for review were liver lesions and pulmonary embolism (PE). Although angiography was considered overall by those surveyed to be the most important diagnostic application of SCT, it was not selected for review for the following reasons: (1) the category was too general (the majority of centres did not specify any anatomical area) and would contain too wide a range of clinical applications; (2) rapid developments in magnetic resonance angiography will rapidly make inaccurate a CT-focused review of angiography; (3) there were other priorities in the Health Technology Assessment Programme specifically addressing angiography; and (4) aspects of angiography (coronary artery disease (CAD)) would be reviewed in association with EBCT.

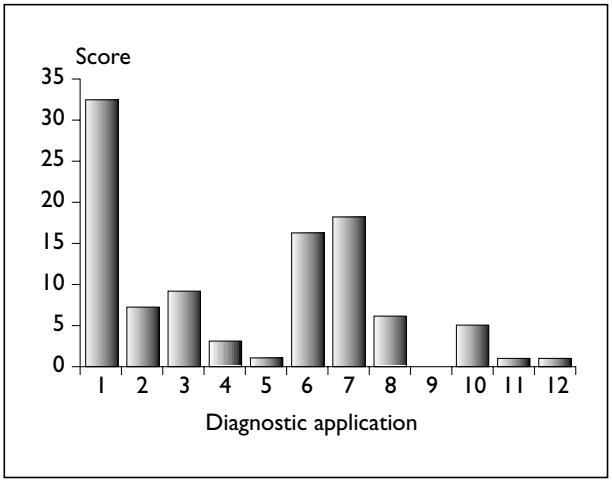


FIGURE 2 *Results of the survey of UK SCT centres (Diagnostic applications: 1, angiography; 2, tumour staging; 3, dental/maxillofacial; 4, gastrointestinal tract; 5, juxtadiaphragmatic lesions; 6, liver lesions; 7, pulmonary embolus; 8, pulmonary nodules; 9, renal calculi; 10, trauma; 11, ear, nose, throat; 12, CT arthrography)*

Liver lesions using SCT

The liver has a dual blood supply, a feature that imaging protocols are able to use to advantage. An overview of imaging of the liver is given by Oliver *et al.*[16] The liver parenchyma is supplied by both the hepatic artery (20–25%) and the portal vein (75–80%). Hypervascular tumours are supplied almost entirely by the hepatic artery, although hypovascular tumours receive minimal flow from either source. Hepatic enhancement after the administration of intravenous contrast agent occurs first via the hepatic artery, about 20–30 s before contrast material arrives from the portal circulation. These two phases are known as the hepatic arterial phase (HAP) and the portal venous phase (PVP). During the HAP, hypervascular tumours are enhanced against a relatively unenhanced parenchymal background. In the PVP, this contrast is reduced because the parenchymal enhancement increases, but hypovascular tumours remain unenhanced, thus becoming more apparent.

The time window for imaging in the HAP is too short for imaging the entire liver with incremental CT, so protocols have concentrated on the PVP or a later phase, where good contrast between liver parenchyma and hypovascular tumours can be obtained. The terminology used in the literature is not consistent, and the later phase is known variously as equilibrium, late or delayed. The advent of SCT and scanners that cool rapidly between acquisitions have allowed the imaging of both HAP and PVP and hence the possibility of visualising hypervascular tumours. The additional HAP would be expected to be an advantage in patients with known or suspected hypervascular neoplasms, when their treatment would be affected by knowledge of the extent of the disease. Such tumours further subdivide into: (1) malignant – hepatocellular carcinoma (HCC), renal cell carcinoma, breast carcinoma, neuroendocrine (islet cell, carcinoid) and melanoma; or (2) benign – focal nodular hyperplasia and hepatic adenoma.

HCC has a poor prognosis associated with its rapid growth; it is also generally associated with liver cirrhosis. The literature on imaging HCC is abundant, and imaging has the potential to provide information that is valuable for patient management. Prognosis is better when there are fewer carcinomas and when they are located in regions of the liver that are amenable to resection, percutaneous ethanol injection or transcatheter arterial chemoembolisation. Segmental location and the extent of the tumour are key criteria for assessing resectability; careful preoperative planning is required to avoid needless surgery.

In this review we shall present the available evidence about the ability of intravenous HAP imaging to visualise lesions additional to those seen in the PVP. It is known that the precise protocol used is important: rate and injection profile; volume and concentration of contrast; and timing of HAP and PVP acquisition. Therefore, for completeness, a review of studies designed to investigate their effects will be presented. Although lesion differentiation and staging are additional areas in which another phase may be hypothesised to be of value, these are not addressed in this review. Alternative imaging investigations for the detection of hypervascular liver tumours are incremental dynamic CT, spiral or incremental CT after iodised oil injection, dynamic MRI, conventional MRI and intraoperative ultrasound (US). We have included a review of studies that compare the performance of SCT acquisitions with one or more or the alternative modalities.

PE using SCT

Acute PE is the occlusion of the pulmonary arterial tree by an embolus, which is usually a large blood clot carried from a deep pelvic or leg vein. A large PE is usually fatal; moderately sized emboli cause haemoptysis, chest pain, breathlessness, hypotension and dizziness. The treatment of a PE is by anticoagulation and, in an emergency, surgery or thrombolysis.

An efficient diagnostic strategy is necessary to assist in the management of suspected acute PE, where risks of morbidity and mortality are associated with treatment as well as the non-treatment of positive cases. False-positive findings are also undesirable because of the dangers of unnecessary anticoagulation. Conventional pulmonary angiography is the standard procedure to guide treatment decisions, but it has associated levels of morbidity and mortality (0.5%[17]), which means that it is often not used when the clinical probability of PE is low. One of the primary non-invasive diagnostic tools, ventilation/perfusion (VP) scintigraphy, is non-specific and can produce several grades of findings. Two of these grades are used confidently for patient management: those described as high probability and normal,[18] respectively, confirm or exclude the diagnosis. However, indeterminate VP findings occur frequently. For example, van Erkel *et al.*[19] have estimated that probably 73% of VP tests produce patterns that are non-diagnostic. Conventional CT is not fast enough for angiography to be performed, but SCT allows volume acquisition in the period of one breath hold[20] and thus has potential as a follow-up test in patients with indeterminate VP findings. Intravenous contrast administration is possible, making it less invasive than selective pulmonary angiography. Alternative strategies have also been proposed, where SCT might itself be the primary diagnostic test or be used together with other modalities as the primary diagnostic test.

However, there is controversy in the literature about the accuracy of SCT for detecting PE. Van Erkel *et al.*[19] reviewed five studies and quoted a sensitivity range of 64–100%. A rigorous review of the literature is required to determine whether more evidence is now available. Only once the true accuracy of SCT for detecting PE is known can its role in the investigation of pulmonary thromboembolism be determined.

Diagnosis and prediction of CAD using EBCT

Like SCT, electron beam systems are favoured for imaging moving structures and dynamic processes, particularly the investigation of cardiovascular diseases.[21] For detecting coronary calcium, a scan slice thickness of 3 mm is used, and up to 40 transverse slices are obtained in one or two breath holds with ECG triggering. Calcium has a high Hounsfield value compared with blood and periarterial fat. A scoring system is used,[22] based on the area of calcified deposits and their Hounsfield number. Although intravascular US, fluoroscopy and X-ray CT may also be used to visualise calcium deposits, EBCT is the only modality that allows the quantification of coronary calcium.

The mortality rate for ischaemic heart disease in England and Wales in 1995 was 2791 and 2239 per million for men and women respectively.[23] This compares with 2795 and 2498 for all malignant neoplasms. In addition, coronary heart disease is one of the priorities of the Health of the Nation programme: "Coronary Heart Disease (CHD), England's biggest single cause of death accounts

for 140,000 deaths per year (25% of all deaths in England in 1994). Between 1990 and 1994 deaths from CHD for under 65s fell by 19.2% and for 65–74s by 12.5%."[24]

In the USA, Ultrafast CT Coronary Artery Disease Risk Assessment Centres are currently being opened at hospital sites across the country, under the auspices of a network called HeartScan Imaging Inc. These centres offer a combination of traditional tests with Ultrafast CT and aim to provide a conclusive, non-invasive assessment of the development of coronary artery calcification. This service is offered to non-symptomatic individuals. It is likely that demand for such assessment centres will arise in the UK, but what is the real evidence about the effective-ness and efficacy of this modality? Currently, it is under assessment in two ways:

- study of its technical ability to demonstrate calcium, with the results fed into separate studies relating calcium to disease development and mortality, or calcium to coronary stenosis[25]
- direct studies seeking the relationship of calcium score to disease development and mortality.

In this review we shall examine the evidence from the second type of study in asymptomatic individuals. These studies correlate the amount of coronary artery calcium (CAC), as detected by EBCT, to subsequent clinical events such as myocardial infarction, angina or coronary artery bypass surgery, and therefore require a period of follow-up to identify the clinical outcomes.

There are two further potential applications of EBCT to CAD, which may be considered as secondary uses for a machine purchased for screening. First is the use of EBCT as an adjunct or alternative to coronary angiography. We shall review studies that compare the diagnoses obtained by the two modalities and seek evidence about the role of EBCT. Should it be used to rule out the presence of disease rather than to diagnose its presence? One of the features to be kept in mind when considering this application is the different manifestations of CAD that are demonstrated by EBCT and angiography. EBCT shows coronary calcium while angiography demonstrates luminal narrowing. The two modalities are not equivalent and, indeed, neither one is a perfect predictor for subsequent cardiac events. Secondly, the potential of EBCT for monitoring disease progression has been subject to debate, with review articles presenting completely opposite views. For example, Viamonte *et al.*[26] state that "Ultrafast CT has a high reproducibility on serial studies", while Wexler *et al.*[27] report that: "The reproducibility studies done to date show that changes in calcium score of as much as 50% may be necessary to be certain that a real change has taken place." Differences in repeated measurements may have multiple causes, including respiratory motion, partial volume effects and observer variability. Wexler *et al.*[27] give a useful overview of the state of the arguments, but their review is not systematic.

Chapter 2

Hypotheses tested in the review

The following questions form the basis of the review. As the review is 'device focused' rather than 'disease focused' some questions encompass a wide range of clinical applications, while, for others, the focus is on specific clinical applications. As a consequence, they are presented in a hierarchy, giving the more general questions first and then becoming more specific in terms of disease or clinical application.

Health economics

- Is there evidence that latest generation CT scanners are cost-effective?
- How cost-effective are latest generation CT scanners compared with the previous generation?

Patient outcome

- Is there evidence that latest generation CT scanners have an impact on patient outcome?

Therapeutic impact

- Is there evidence that latest generation CT scanners have an impact on patient management?

Diagnostic impact

- Can latest generation CT scanners replace existing investigative techniques?

Diagnostic performance

- In which clinical conditions and disease groups are latest generation CT scanners of potential value?
- What are the diagnostic benefits of latest generation CT scanners? Can these be expressed in terms of sensitivity and specificity? The clinical applications that have been addressed are listed below.

Liver lesions using SCT

- Are more lesions (of all types) seen when using SCT protocols that include the HAP than in those with just the PVP?
- What is the detection accuracy of the HAP for HCC?
- Is there any evidence directly comparing the SCT HAP detection accuracy with other modalities?
- With reference to injection rate and profile, delay time and contrast dose, is there an optimal protocol for SCT liver investigations? If so – what is it?
- What is the effect on patient management of the detection of lesions compared with no lesions detected?

PE using SCT

- What evidence is there about the accuracy of SCT in the detection of PE?
- Can the diagnostic strategy for the detection of PE be improved by including SCT?

CAD using EBCT

- Can EBCT predict CAD in asymptomatic individuals?
- For symptomatic patients (chest pain), what is the diagnostic accuracy of EBCT in CAD compared with the gold standard of coronary angiography?
- Can EBCT be used to track the progression of coronary atherosclerosis?
- Is there any evidence directly comparing the predictive performance of EBCT with other tests?

Other hypotheses tested

Although this review does not deal with the evaluation of the technical aspects of latest generation CT scanners, one device-related technical aspect has been addressed, that of radiation dose. It is an important factor that will have an influence at the patient outcome and health economics levels of the hierarchy.

- What evidence is there about the radiation dose associated with latest generation CT scanners?

- Is there any evidence or published work relating radiation dose to the clinical use of latest generation CT?

In addition to reviewing evidence to answer questions specifically associated with latest generation CT scanners, this systematic review offered the opportunity to address more generic questions concerning the methodology of the technological assessment of diagnostic devices.

- Can we develop a methodology for systematic reviews for diagnostic devices that are continually evolving?
- Is it possible to combine sensitivity and specificity data from different publications?

Chapter 3
Review methods

General methodology

A multidisciplinary review team was assembled; its composition was designed to:

- ensure a broad spread of relevant expertise
- minimise the potential for bias in the review
- facilitate the dissemination of both review methodology and review results among several professional groups.

The panel comprised the authors and an external member, who is an opinion leader in the field of radiology. The professions represented were medical physics, radiology, radiography, health economics and public health medicine.

The methodology is broadly based on that recommended in Centre for Reviews and Dissemination Report 4,[28] but it was necessary to adapt the approach for this 'device driven' review of a medical imaging modality.

This review is organised in sections corresponding with the levels of the evaluative framework:[6–8]

- health economics
- patient outcome
- therapeutic impact
- diagnostic impact
- diagnostic performance (or accuracy)
- radiation dose.

Search strategy

The search strategy was primarily technology based and the resources described in this section were all searched for information on latest generation CT scanners. For information regarding the specific levels of the evaluative framework, more precise searches were conducted using MEDLINE, Bath Information and Data Services and the Institute of Scientific Information (BIDS-ISI). To address the dosimetry implications of latest generation CT, an additional search was performed using dose keywords in the title or abstract from studies within the database of already identified articles. In addition, any information on dose was highlighted throughout the reviewing process. All search strategies are shown in appendix 1.

Electronic databases

The majority of publications associated with medical imaging are available electronically. The following databases were searched and the search strategies are given in appendix 1:

- MEDLINE
- BIDS-ISI
- EMBASE
- Cochrane Library
- Inside Information Plus (British Library)[29]
- FirstSearch Online Computer Library Center (OCLC).[30]

Unlike MEDLINE, BIDS-ISI does not classify articles into subject categories, but it does provide similar Boolean and text word capabilities. The BIDS-ISI archive extends from 1981 to the present day. It is updated daily, which is an advantage over MEDLINE. BIDS-ISI also includes selected conference proceedings and abstracts, a service not supplied by MEDLINE. There is substantial overlap between these two databases. In addition, a small proportion of articles are unique to them individually. A comprehensive comparison of MEDLINE and BIDS-ISI search strategies demonstrates subtle differences between Boolean commands and also the incompatibility of specialised search commands of these systems. Hence, separate search strategies were compiled for these two databases. Both MEDLINE and BIDS-ISI were searched from 1981 to the end of 1996. As there are sometimes delays before MEDLINE updates with new publications, the search strategy was re-run in October 1997 to ensure all references were up to date. The EMBASE search, performed by a library professional, was designed to have higher precision and therefore lower recall, to limit the number of inappropriate retrievals.

The last two resources listed above became available in 1997, during the period of this review; they facilitated access to otherwise inaccessible journals. The first, Inside Information Plus, supplied by the British Library, enables access, searching and ordering of a large selection of their archive. The service covers 250,000 journals, of which 20,000 can be searched down to article title level with the

use of keywords, together with 16,000 conference proceedings. The archive extends back only as far as 1993, but the service is updated within 72 hours of receipt of new material. The second service is provided by the OCLC, which is a non-profit computer service and research organisation whose network and services link more than 25,000 libraries in the USA, and 63 countries and territories. Using a service called FirstSearch, more than 60 databases in 14 topic areas can be searched by using keywords. These databases include: WorldCat, a merged electronic catalogue of libraries around the world; ArticleFirst, a catalogue of individual articles; ContentsFirst, a catalogue of journal contents divided into volume and issue; NetFirst, a catalogue of Internet-accessible resources; and Proceedings-First, a catalogue of conference proceedings. All these databases were searched by title and subject using the technology-based keywords included in the MEDLINE strategy.

Handsearching

Journals that were cited by one of the main electronic databases (MEDLINE or BIDS-ISI) were not handsearched. Because high-recall search criteria were used and additional resources were searched, the impact of not undertaking this extensive task is considered to be negligible. In addition, for confirmation, some health economics journals were handsearched. Uncited journals were identified from the reference lists of articles from cited journals, from the ISI citations lists, by browsing library catalogues, and from Internet web sites. The reference lists of all retrieved articles were handsearched to identify any additional studies. Textbooks were not searched.

The electronic availability of journals that are not on MEDLINE or BIDS-ISI is listed in *Table 1*.

Many of the journals initially identified as requiring handsearching are included in the

TABLE 1 *Summary of listing in electronic databases of journals not cited in MEDLINE or BIDS-ISI*

Journals not on MEDLINE or BIDS-ISI	Searched electronically	Cited by ISI	Inside Information Plus	FirstSearch
Acta Chirurgica Austriaca	✔	✗	✔	✗
Advanced Imaging	✔	✗	✔	✔
Annual of Cardiac Surgery	✗	✗	✗	✗
Applied Radiology	✔	✗	✔	✔
Asian Journal of Surgery	✔	✗	✔	✗
Asian Medical Journal	✔	✗	✔	✔
Cardiology Clinics of North America	✗	✗	✗	✗
Chirurgia	✔	✗	✔	✗
Clinical MRI	✔	✗	✔	✗
Contributions to Oncology	✔	✔	✗	✔
Current Oncology	✗	✗	✗	✗
Diagnostica	✔	✔	✔	✗
Diagnostic Imaging	✔	✗	✔	✔
Emergency Radiology	✔	✗	✔	✗
Evidence Based Medicine	✔	✗	✔	✗
Indian Journal of Radiology and Imaging	✔	✗	✔	✗
Journal of the Japan Society for Cancer Therapy	✔	✔	✔	✗
Journal of Medical Imaging	✗	✗	✗	✗
Surgical Research Communications	✔	✗	✔	✔
World Congress – International College of Surgeons	✔	✗	✔	✗

two new services, Inside Information Plus and FirstSearch; thus the handsearching task was substantially reduced. The four journals that were not possible to search electronically were excluded for the following reasons:

- *Journal of Medical Imaging* was incorporated into the *European Journal of Radiology* for the period 1987–1989; this journal was searched electronically on MEDLINE
- *Annual of Cardiac Surgery and Current Oncology* contained summary reviews
- *Cardiology Clinics of North America* was not held at the British Library and was not catalogued by any database, including the international periodicals directory, Ulrich.

Contacting authors, academic centres and manufacturers

We chose not to write to authors of conference abstracts to ask if the work described had since been published because of a very low response rate, which produced no unknown studies, to a mailing performed in another systematic review.[31] Details of ongoing studies were requested from UK SCT centres, as described in the 'Diagnostic performance' section of this chapter. Six mainstream manufacturers of CT equipment were identified and asked to provide information on publications and journals.

Grey literature

The database of grey literature supplied by the British Library, the System for Information on Grey Literature, was searched using technology-based keywords.

Inclusion criteria

Study selection was a three-stage process. First, for all levels of assessment, preliminary inclusion criteria were applied to the returns of the electronic searches:

- published before January 1997
- not an abstract
- not a review article
- English language
- not a case report
- not an editorial
- not a letter.

Secondly, simple exclusion criteria were applied in order to ensure the applicability and utility of the studies. These excluded non-human studies, studies with ten or fewer patients, and inappropriate studies that had been retrieved from a keyword used in a different context from that intended. For diagnostic performance and diagnostic impact studies, subject-specific inclusion criteria were set to maintain relevance to the chosen topics. Full details are given in the following sections containing the words '– specific inclusion criteria'.

The final set of criteria assessed the relevance and validity of the articles retrieved and was used to select those suitable for inclusion in the review. These criteria depended on the application and are described in the following sections containing the words '– assessment of relevance and validity of primary studies'.

Health economics

Specific search strategy

A broad search strategy was adopted, using search terms for economics and SCT or EBCT in any clinical application. As well as the medical subject heading (MeSH) category 'economics', several relevant economic text word indicators were used individually to ensure a search with high sensitivity. Other terms frequently (but not exclusively) used in economic studies – 'benefit', 'impact', 'management outcome' and 'utility' – were limited to a combination of any two, to balance between retrieving all relevant studies and minimising the identification of inappropriate articles. The search terms used in MEDLINE and BIDS-ISI are given in appendix 1, together with those for EMBASE. No search specific to economics was performed using the remaining databases.

Handsearching was undertaken of the bibliographies of articles identified in the electronic search, selected health economics articles identified in the electronic search, and selected health economics journals. The following journals were handsearched:

- *International Journal of Technology Assessment in Health Care*
- *Health Economics*
- *Health Policy*
- *Social Science in Medicine.*

Abstracts from the 1995, 1996 and 1997 International Society of Technology Assessment in Health Care conferences were also handsearched.

Specific inclusion criteria

The abstracts of retrieved articles were read and a decision was made about the relevance of each study in terms of the applicability to latest generation CT scanners, and the fulfilment of the preliminary criteria. Neither the anatomical area nor the clinical application was used as an exclusion criterion. Full copies of qualifying articles were acquired and rechecked against the preliminary criteria. Further exclusions were made as necessary.

The decision of whether or not to include an article was based solely on economic information being reported.

Assessment of relevance and validity of primary studies

Economic evaluations of healthcare technologies have been undertaken regularly for over 25 years. Agreement on the most appropriate methods has been facilitated by practical experience and the refinement of economic techniques. As a result, there is no shortage of guidelines and checklists to assist the reader of economic studies. Williams[32] sets out the fundamentals, which have subsequently been elaborated by the Department of Biostatistics, McMaster University[33] and Drummond and colleagues.[34] In 1996, the *BMJ* published a set of guidelines for use by reviewers of economic submissions to the Journal.[12] Condensed lists of key factors have been used by some authors in empirical studies of the quality of economic evaluations found in the clinical literature.[35,36] Adaptations of these published guidelines have been used in other recent studies of the economic evaluation of diagnostic imaging.[37]

The initial intention was to do something similar in this review. However, it became apparent early on in the project that the quality of the economic analyses in the studies located was so poor that the use of a long checklist to assess quality was redundant. The proposed checklist, which was designed but not used, is shown in appendix 2.

To classify studies into those that:

- had adequate economic analyses
- had poor economic analyses
- could not legitimately be called economic analyses,

the following four criteria were identified as sufficient:

- Was the type of economic analysis correctly chosen and designed?
- Was the outcome indicator appropriate?
- Was the cost analysis correctly conducted?
- Was sensitivity analysis carried out?

Data extraction

Descriptive summaries were written.

Data synthesis

Neither quantitative nor qualitative data synthesis was applicable at the health economics level, owing to the small numbers of studies available.

Patient outcome

Specific search strategy

The additional search strategy for patient outcome and therapeutic impact is shown in appendix 1. This specific search was performed only on MEDLINE and BIDS-ISI. A change in patient outcome may, or may not, result from the cumulative changes in diagnostic and therapeutic impact. Studies not following strict methodologies to evaluate patient outcome, but supplying limited information regarding patient survival, were identified but not assessed.

Specific inclusion criteria

The abstracts of the articles were read and a decision was made about the relevance of each study in terms of the applicability to latest generation CT scanners, the fulfilment of the preliminary criteria, and the following exclusion criteria:

- keyword used in a different context from that intended in our search
- ten or fewer patients
- non-human study.

Neither the anatomical area nor the clinical application was used as an exclusion criterion. Full copies of qualifying articles were acquired and rechecked against the criteria previously described, with further exclusions made as necessary.

The decision about whether or not to include an article was based on the following simple criteria, assessed from reading the full article. Because the level of information provided was poor, these basic criteria were all that were used for patient outcome studies:

- original data reported (not just qualitative discussion)
- comparative study with and without latest generation CT
- link between the use of latest generation CT scanners and patient outcome.

Assessment of relevance and validity of primary studies

The imaging literature reviewed was lacking in good studies that assessed outcome. As a consequence, a strict checklist of criteria was not applied and all studies that promised any outcome data were included. The criteria for a valid study designed specifically to evaluate patient outcome have been discussed in chapter 1. No formal assessment was performed because there was a paucity of data. The limitations of identified studies are detailed in chapter 8.

Data extraction and synthesis

Descriptive summaries were written but the differing study designs did not allow any qualitative or quantitative data synthesis to be performed.

Therapeutic impact

Specific search strategy

The additional search strategy for patient outcome and therapeutic impact is shown in appendix 1. This strategy was performed only on MEDLINE and BIDS-ISI.

Specific inclusion criteria

The article abstracts were read and a decision made about the relevance of each study in terms of the applicability to latest generation CT, the fulfilment of the preliminary criteria, and the following exclusion criteria:

- keyword used in a different context from that intended in our search
- ten or fewer patients
- non-human study.

Neither the anatomical area nor the clinical application was used as an exclusion criterion. Full copies of qualifying articles were acquired and rechecked against the criteria previously described, with further exclusions made as necessary. The decision of whether or not to include an article was based on the following simple criteria, assessed from reading the full article. Because the level of information provided was poor, these basic criteria were all that were used for therapeutic impact studies:

- original data reported (not just qualitative discussion)
- comparative study with and without latest generation CT
- link between the use of latest generation CT and changes in therapeutic decisions or confidence in those decisions.

Assessment of relevance and validity of primary studies

Criteria for a valid study designed specifically to evaluate therapeutic impact have been discussed in chapter 1. No formal assessment was performed because there was a paucity of data. The limitations of studies identified are detailed in chapter 8.

Data extraction and synthesis

Descriptive summaries were written, but the variety of study designs did not allow any data synthesis at the therapeutic impact level.

Diagnostic impact

Specific search strategy

No special search strategy was used; the returns of the searches at the patient outcome, therapeutic impact and diagnostic performance searches were used.

Specific inclusion criteria

Studies that fulfilled the major subject-specific criteria described in the next section on diagnostic performance (p. 16) were reviewed for studies comparing modalities. The preliminary criteria and the following exclusion criteria were applied:

- ten or fewer patients for any modality
- non-human study
- duplicate patient data set (only the most recent report for a given patient group was included)
- independent reference standard not used.

Assessment of relevance and validity of primary studies

The criteria for a valid study designed specifically to evaluate staging impact have been discussed in chapter 1. No formal assessment was performed owing to the paucity of data. For studies that compared latest generation CT with other modalities, the validity was assessed by following the methodology described in the next section for studies of diagnostic performance.

Data extraction and synthesis

Descriptive summaries were written for studies designed to address diagnostic impact, but no data synthesis was performed owing to the paucity of data. For comparative studies, similar data extraction and synthesis to that for diagnostic performance studies (next section) was planned if sufficient studies comparing the same modality were identified.

Diagnostic performance

The performance of a diagnostic test can be measured by comparing its results with the truth. In the field of diagnostic imaging, the best approximation of the truth is that provided by a gold standard test. Comparative statistics such as sensitivity, specificity and accuracy can then be used as summary estimates of the test's performance. The limitations and threats to the validity of such techniques are well recognised and will be described in context.

The next subsection describes more fully the process undertaken to determine the specific clinical applications to review for SCT and EBCT. After this preliminary description, the remaining subsections outline the processes involved in searching for diagnostic performance studies, identifying relevant studies, assessing study validity, extracting data on diagnostic performance and, finally, methods of data synthesis.

Determination of clinical applications for SCT and EBCT

SCT

We contacted opinion leaders at 19 centres in the UK that had access to SCT machines. A choice of the nine diagnostic applications of SCT shown in *Table 2* was presented and they were asked to indicate and rank (1–3) the three they felt were the most important.

The response was excellent, with all 19 forms returned. There were 16 completed forms and three uncompleted, two because no SCT machine was available at the location and one suggesting an alternative contact with more experience in SCT. The results of the 16 replies are shown in *Figure 2*. Each diagnostic application received a score of 3 if it was ranked as the highest importance, and 1 as the lowest importance. A score of zero was assigned for those not chosen. Four centres specified three more of their own diagnostic applications. Of the nine originally specified, only one was not selected at all, namely, renal calculi.

TABLE 2 *Part of the questionnaire sent to UK opinion leaders in SCT*

Diagnostic application	Rank
Angiography (if specific anatomical area, please specify)	
Tumour staging (if specific anatomical area, please specify)	
Dental/maxillofacial	
Gastrointestinal tract (if specific anatomical area, please specify)	
Juxtadiaphragmatic lesions	
Liver lesions	
Pulmonary embolus	
Pulmonary nodules	
Renal calculi	
Other: please specify	
Other: please specify	
Other: please specify	

The diagnostic applications that were selected for review were liver lesions and PE. The reasons were given in chapter 1.

To maximise the available information, the questionnaire had an additional item about ongoing trials of SCT. Nine centres supplied information detailing 11 studies in progress or due to commence. Of these, two were based on the liver and one on PE. This information is included in chapter 8 and will be helpful for future updating of the results of this review.

After identifying these two major clinical applications of SCT, further classification of the subject areas for review was determined from preliminary searches of the topics and consensus decisions from the multidisciplinary panel. The chosen subjects were, for SCT for liver lesions:

- comparison of HAP and PVP for lesion (any type) detection
- accuracy of HAP in detecting HCC
- comparison of a PVP protocol for lesion detection or liver enhancement.

Those for SCT for PE were:

- detection accuracy of acute PE by SCT in place of VP

- detection accuracy of acute PE by SCT together with VP
- detection accuracy of acute PE by SCT with other modalities.

EBCT

The major topic of CAD was selected as part of the remit of the review. The following specific categories within this topic were identified:

- prediction of CAD in asymptomatic individuals
- diagnostic accuracy for CAD in symptomatic patients
- reproducibility or observer variation for tracking disease progression.

Specific search strategy

The full electronic search strategies shown in appendix 1 were used for SCT and EBCT. They were not restricted to the specific clinical applications. The identification of clinical application was achieved by reading the abstracts, as described in the next subsection.

Specific inclusion criteria

The preliminary inclusion criteria were applied. The remaining abstracts and titles were assessed against the inclusion criteria shown in *Table 3*.

In each major topic of SCT and EBCT, additional criteria were set for the specific applications. These are summarised in *Table 4*.

After these subject-specific criteria, the following exclusion criteria were applied:

- ten or fewer patients
- non-human study
- duplicate patient data set (only the most recent report for a given patient group was included).

Full copies of qualifying articles were acquired and rechecked against all the criteria previously described, with further exclusions made as necessary. The remaining studies were included in the qualitative review. For inclusion in a quantitative analysis, the additional criteria of *Table 5* were applied.

Assessment of relevance and validity of primary studies

The Centre for Reviews and Dissemination (Report 4)[28] recommends grading primary studies into a hierarchy according to the study design. Level I includes well-designed randomised controlled trials; level II includes both prospective and retrospective controlled trials; while level III covers comparisons that are lacking controls. Level IV is opinion-based evidence. Because controlled studies were not found in our topic area, meaning that all our evidence falls in level III, this hierarchy proved to be inapplicable to this review. Instead, it was decided to begin our assessment of validity at a lower level, by determining the presence in the study design of features that could lead to bias that was likely to

TABLE 3 *Major application-specific inclusion criteria for diagnostic performance studies*

Inclusion criterion	SCT – liver	SCT – PE	EBCT
Anatomical location	Liver	Lung	Heart
Type of disease	Cancer	Acute PE	CAD

TABLE 4 *Application-specific inclusion criteria for diagnostic performance studies*

Major topic	Specific application	Criterion
SCT – liver	HAP vs PVP	Information on both phases provided
	HAP for HCC	Reference standard used
	Protocol	Comparative study ensuring only one variable between groups
SCT – PE	In place of VP	Reference standard used
	With VP	Reference standard used
	With other modalities	Comparative study or reference standard used
EBCT	Prediction	Patient follow-up > 1 year
	Diagnosis of CAD	Reference standard used and calcium score reported
	Reproducibility	Time elapsed between scans < 1 week

TABLE 5 *Inclusion criteria for quantitative summary receiver operator characteristic (SROC) meta-analysis of diagnostic performance studies*

Inclusion criterion	Value
Adequate gold standard	Pathology/histology for liver Angiography for PE Angiography for CAD
Sufficient raw data presented	To enable completion of 2 x 2 contingency table
Comparable definitions of dichotomy	To enable similar studies to be combined

TABLE 6 *Potential biases in diagnostic imaging studies (modified from reference 15)*

Subjects			Study			Interpretation
Patient selection			***Application of the gold standard***			***Independence of interpretations***
Referral bias	*Patient filtering*	*Patient cohort*	Verification bias			Diagnostic review
Centripetal	Diagnostic safety	Spectrum	Work-up bias			Test review
Popularity	Co-intervention	Population	Incorporation bias			Comparator review
Diagnostic access						Clinical review
			Measurement of results			
			Disease progression	*Withdrawal bias*	*Observer variability*	
				Indeterminate results	Intrinsic interobserver	
				Loss to follow-up	Extrinsic interobserver	
					Intraobserver	
Main effect external validity			**Main effect internal validity**			**Main effect internal validity**

threaten validity. We drew up a list of 20 potential biases, which are shown in *Table 6* and described in full by Kelly *et al.*[15]

Identifying the presence of bias

A checklist approach was required, but we found that none of those published at the time we began our review was suited to the application. Those designed for randomised controlled trials were inapplicable; those designed for observational studies were best suited to controlled trials of treatment; and even those for diagnostic tests[38–40] were not as generic as we wished. For example, in medical imaging, a test may be used for purposes other than to differentiate between affected and disease-free individuals. Tests may be used for staging disease, as part of a diagnostic work-up, or to guide other procedures. We required a checklist that was generic enough to be adapted quickly to suit both pure diagnostic applications and other studies. In addition, those features of the study conduct that might vary between studies must be noted to allow proper comparison of studies. We have chosen to call this category 'factors', and our checklist has a separate section to note information relating to the equipment used and the imaging protocol.

The checklist is a two-part document: the questions that are generic to all applications; and essential guidelines that are specific to the clinical question. A generic bias checklist with guidelines, and factor checklists for SCT of liver

lesions, SCT for PE, and EBCT for CAD, are shown in appendix 3. The bias checklist comprises 30 questions divided into four major sections; the first section covers the focus and basic details of the article and the remaining three cover biases related to patient selection, biases related to study conduct, and independence of interpretation biases. The checklist was designed to assess individual study quality by containing specific questions applicable to each of the potential biases, while maintaining a broad applicability over all diagnostic performance studies. In order that the answers to the questions should be reproducible and objective, very specific guidelines are required. These guidelines may require slight modification for different clinical applications, which was the case for the three topics reviewed. For the factor checklists, variation occurs as a result of the inherent differences between the technologies and/or applications.

Our checklist is compatible with the suggestions of the Cochrane Methods Working Group on Systematic Review of Screening and Diagnostic Tests.[41] It covers very similar points, but is presented as a series of questions.

Assessment of the checklist's interobserver reliability was performed on two separate occasions; this is described elsewhere.[31] Changes were made after the first trial. The checklist was shown to be sufficiently reproducible and objective for our purposes, and also reasonably simple to complete.

Ranking study validity

After completing the checklist, we had hoped to be able to rank the biases in order of significance (in a manner similar to that described by Mulrow *et al.*[38]) and develop a numerical scoring scheme that would allow the objective ranking of studies by validity.[42] The investigation of this approach has been reported in another *Health Technology Assessment* review.[31] This approach was abandoned because no consensus could be reached by our review team on the relative importance of the biases. This meant that any scoring system we produced could not be objective because it would be based on a controversial choice of rankings. Even if an unweighted combination is considered there are difficulties. Studies in a given subject tend to have properties in common, perhaps dictated by the clinical application area, which means that they share common faults in study conduct, data interpretation or patient selection, and that they will not be differentiated in such a scheme.

A further difficulty arose because of a widespread lack of reporting of study design in the medical imaging literature. In common with most authors[43] we chose to rate a study as having a design that could cause a particular bias if the information required to determine whether or not a design feature existed was not given. A very high proportion of studies fell into this category and it was clear that the final review would exclude potentially valid results because of this. Instead, all studies meeting the inclusion criteria were incorporated in a statistical analysis designed to determine if the study results were related to the likely presence of one or more biases.

Data extraction

For SCT of liver lesions and EBCT, data were extracted from the studies by the main reviewer. The PE SCT data extraction was performed by an additional panel member together with the main reviewer. The factor and bias checklists (appendix 3) were completed and the diagnostic performance results were extracted, where relevant. *Figure 3* shows the results table that was completed for SCT PE studies and EBCT, which were the only topics for which numerical results could be extracted. For review topics for which quantitative results were not available or reliable,

	Gold standard		
Test	Positive	Negative	Total
Positive	TP	FP	TP + FP
Negative	FN	TN	FN + TN
Total	TP + FN	FP + TN	N

$$Sensitivity = \frac{TP}{TP + FN} \qquad Specificity = \frac{TN}{TN + FP}$$

$$PPV = \frac{TP}{TP + FP} \qquad NPV = \frac{TN}{TN + FN}$$

$$Accuracy = \frac{TP + TN}{N} \qquad OR = \frac{TPR/(1 - TPR)}{FPR/(1 - FPR)}$$

FIGURE 3 *2 x 2 Contingency table and equations for expressing staging performance (TP = true-positive, TN = true-negative, FP = false-positive, FN = false-negative, N = TP + TN + FP + FN, PPV = positive-predictive value, NPV = negative-predictive value, OR = odds ratio, TPR = true-positive rate = sensitivity, FPR = false-positive rate = 1 – specificity)*

descriptive summaries were written. For example, in the review of SCT of liver lesions, owing to the lack of confirmatory data for negative diagnoses, only incomplete or pseudo-accuracy results were available.

In order to obtain the results to complete the 2×2 table, it is necessary to define the categories 'positive' and 'negative' for both the diagnostic test and the gold standard. For the PE review this division (dichotomy) is simple. A positive result is one where PE is considered present according to the test in question, be it the gold standard or another investigation. Similarly, a negative result is one where PE is diagnosed as absent.

For EBCT a more detailed description of what is positive and negative is required. First, for the gold standard of angiography, CAD is diagnosed from the narrowing of the coronary arteries. By defining different thresholds of percentage diameter or the area of luminal stenosis, the dichotomy between negative and positive is classified. For example, a possible dichotomy is for a positive diagnosis for CAD to correspond with any luminal narrowing (> 0% stenosis) and negative to correspond to no narrowing (0% stenosis).

Secondly for EBCT, the diagnosis of CAD is generated from the amount of CAC present. A scoring system proposed by Agatston *et al.*[22] is generally used to quantify the amount of CAC. Again, various cut-offs of this CAC score can be used to form the division between positive and negative. A common dichotomy is for a score of zero to represent a negative diagnosis for CAD and a score of greater than zero to represent a positive diagnosis.

As a variety of thresholds can be arbitrarily chosen, both for the EBCT test and the gold standard, these values were recorded. Accuracy results calculated using the expressions in *Figure 3* are dependent on the value of these thresholds, and therefore only studies using identical thresholds are suitable for direct comparison.

Data synthesis

SCT – PE

The results from the SCT PE review were divided into two categories, depending on the coverage of the scan:

- central PE
- peripheral PE.

The results from these two categories are inherently different and so cannot be combined. Each category was analysed separately.

The results of each primary study were expressed using the summary statistics: sensitivity, specificity, positive predictive value (PPV), negative predictive value (NPV), accuracy and the odds ratio (OR). The results for each study are shown in appendix 4. Visualisation of these results was aided by the use of simple receiver operator characteristic (ROC) scatter plots. Because so few studies were retrieved, further analysis was inappropriate.

EBCT – CAD

Data synthesis was applied to results investigating the diagnostic performance of EBCT for CAD. As emphasised in the discussion of EBCT dichotomy, only studies that used the same explicit thresholds for defining positive and negative for both the gold standard and the EBCT scan can be directly compared. The results of such studies can still differ owing to intrinsic operator dependence. A technique based on ROC methodology[44] was used to combine the results from several studies.

The results of each primary study were expressed using the summary statistics: sensitivity, specificity, PPV, NPV, accuracy and the OR. These are shown for each study in appendix 4. Although the calcium score is apparently an objective measure, observer judgement is involved in region definition, so it was possible that all of these statistics were observer dependent, meaning that there was effectively a different threshold between positive and negative used by observers in each study. Thus, a range of values is expected for each statistic. This is analogous to ROC methodology,[44] where a range of different thresholds is used to plot a curve representing the performance of the test, independent of a preselected threshold. It is then possible to select a threshold between positive and negative that gives the desired balance of sensitivity and specificity. In this review, the results of the independent trials were combined by using the methodologies developed by Moses *et al.*[45] and Irwig *et al.*,[46,47] which expand on the principles of ROC analysis. The overall procedure is outlined in *Table 7*.

Stage 1: ROC scatter plot to visualise range of results from the primary studies

An ROC scatter plot was made to illustrate the range of results for the tests. It was used for visualisation and to show if any of the individual

TABLE 7 *Summary of SROC meta-analysis for synthesis of results from diagnostic performance studies*

Stage	Plot	Reason
Stage 1	ROC scatter plot	To visualise the range of results from the primary studies Variations in reported sensitivity and specificity are assumed to result from the use of different thresholds for defining positivity
Stage 2	Linear SROC with axes D and S, where D and S are as defined in equation 1	Straight line fitted to estimate a best fit to the data under logistic transformation, representing an underlying common test performance
Stage 3	SROC curve with conventional axes TPR and FPR	To present the combination of results from primary studies as a single SROC curve

results lay outside the area of decision making, defined as sensitivity and specificity both greater than 50%.[45]

Stage 2: SROC curve to estimate a best fit to the data under logistic transformation

An SROC was fitted according to the linear model shown in equation 1, in order to represent the underlying common test performance in the absence of possible relations between the results and the threshold used within a study for classifying results as positive. To avoid problems from missing points caused by zero cells of the 2×2 contingency table (i.e. a test with either zero or 100% sensitivity or specificity), the value in each cell was increased by 0.5.[45] Two models[45] were used to fit the line, the equally weighted least squares (EWLS) method and a robust-resistant (RR) method.

The intercept, A, of the model is the estimated ln(OR) when sensitivity equals specificity (S = 0). The gradient, B, provides an estimate of the extent to which the ln(OR) is dependent upon the threshold used. If B is zero, ln(OR) is independent of threshold and test accuracy may be summarised by a common OR, given by the intercept A.

$$D = A + BS \qquad \text{equation 1}$$

where:
$D = \text{logit}(TPR) - \text{logit}(FPR)$
$S = \text{logit}(TPR) + \text{logit}(FPR)$
and
$\text{logit}(x) = \ln\{x/(1-x)\}$

Stage 3: SROC curve to present combination of results from primary studies

An SROC curve, with conventional TPR and FPR axes, was plotted to summarise the combined results. Equation 2 was used to convert back to the conventional axes, with substitution of the gradient and intercept values calculated from the model of Stage 2, providing the gradient calculated was non-zero.

$$TPR = \left[1 + e^{-A/(1-B)} \left(\frac{1 - FPR}{FPR} \right)^{\frac{(1+B)}{(1-B)}} \right]^{-1} \qquad \text{equation 2}$$

Note that TPR = sensitivity, and FPR = (1 – specificity).

Several parameters can be given to summarise this SROC curve, including:

- the point on the curve where sensitivity is equal to specificity (denoted as Q*)
- the area under the curve
- the TPR read at the mean FPR.

The appropriate summary value depends on the intended application of the diagnostic test. For example, for a test where the correct diagnosis of negative patients is equally as important as the correct diagnosis of positive patients, the point on the curve where sensitivity equals specificity is the best summary value. In the case of EBCT, such a balance between sensitivity and specificity is not the most appropriate parameter. In the role of EBCT for asymptomatic patients, as a screening application, a high sensitivity at the cost of specificity is the optimal threshold; whereas EBCT for symptomatic patients would be optimised with a high specificity at the cost of sensitivity. Therefore, the TPR at the mean FPR was chosen to reflect this choice.

Differences between studies

Further analysis was possible only for EBCT.

Regression techniques were used to analyse the influence of differences between the studies on the summary results. The covariates for the analysis were the biases and factors described earlier in this chapter. *Table 8* lists those included. Not all biases initially identified are included in *Table 8* because

TABLE 8 *Factors and biases included in the multivariate analysis of the diagnostic performance of EBCT for asymptomatic CAD*

Biases	Factors
Disease progression bias	No. slices
Blinding biases	No. pixels
	No. patients

the studies had similar gaps in information and also some bias risks were present (or absent) in all articles. The blinding biases represent the four biases labelled under 'Independence of interpretations' in *Table 6*. If any attempt to perform blinding was reported, this combined bias was classified as absent.

For this part of the analysis, the EWLS method was used and the factors or biases were analysed by incorporating them as a multivariate extension into equation 1.[45] Using this methodology, the SROC curve is divided into two separate plots for each variable assessed; *Figure 4* illustrates the principles. The gradients of each of the two separate plots are made equal; the fit of this gradient is dependent on the EWLS plots of each set. If the variable is found to be significant in the regression analysis, then each of the two sets will have a significantly different intercept. Compared with the intercept of the total data set, one set will have a higher intercept and the other a lower intercept. Thus the range between the two intercepts represents the likely variance of the overall result.

If no factors or biases were found to be significant when combined together in the multivariate regression analysis, then regression was performed individually for each variable.

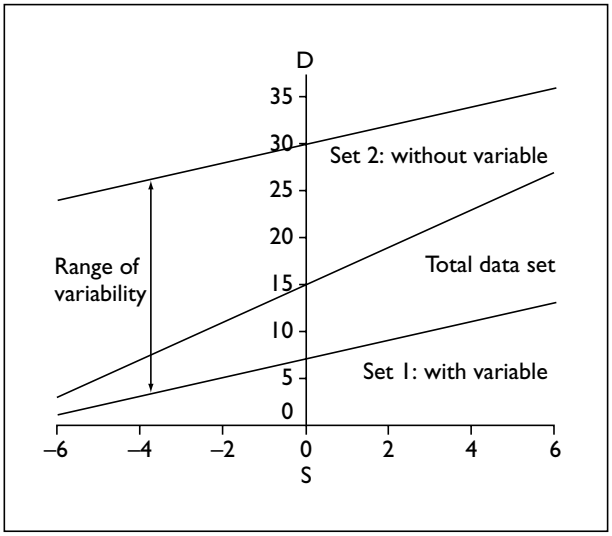


FIGURE 4 *Illustration of regression analysis on an SROC curve to compare two subgroups*

Radiation dose

Specific search strategy

To address the implications associated with radiation doses from latest generation CT scanners, all studies within the databases built up from other searches were searched using keywords within the title or abstract. The keywords used are given in appendix 1.

Specific inclusion criteria

The preliminary inclusion criteria and the following exclusion criteria were applied:

- keyword used in a different context from that intended in our search
- duplicate report (only the most recent report for a given patient group was included)
- therapeutic dose rather than radiation dose
- contrast media dose rather than radiation dose.

Non-human and phantom studies were included in this case.

The remaining abstracts and titles were assessed against the following inclusion criteria:

- comparative study of CT versus other modality – both with dose information
- comparative study of CT protocols – both with dose information.

Assessment of relevance and validity of primary studies

The aim of these comparative criteria was to ensure that the study provided valuable and reliable information on a stand-alone basis. This is important because any studies that do not provide this information require external comparison to place them in context. This external comparison can be unreliable owing to confounding factors between the studies, such as differences in the dosimetry or technological protocols. In addition, to ensure a comprehensive comparison of the dose between comparative modalities it would be important to search extensively for both modalities, an undertaking outside the remit of this review.

No other assessment of validity was performed.

Data extraction and synthesis

Information was recorded and presented in tabular form to enable comparison. No quantitative data synthesis was performed.

Conclusion

In this chapter we have described the methodology of our review, dividing the studies reviewed into five categories. In the next chapter, details of those studies that satisfied the inclusion criteria are presented, as outlined in the 'data extraction' sections of this chapter. *Table 9* summarises where the results of applying the various parts of the methodology are presented in the review.

TABLE 9 *Locations in the monograph of the results of the review*

Description	Results chapter/s
Search strategy	4
Inclusion criteria	4 included studies 5 excluded studies
Assessment of relevance and validity of primary studies	7
Data extraction	4
Data synthesis	6
Differences between studies	7

Chapter 4

Details of studies included in the review

Detailed analysis of search methodology

Health economics

From the MEDLINE and BIDS-ISI search strategies combined, 85 candidate articles were retrieved for SCT. Of these, 26 were not applicable to SCT and 23 were abstracts from conference proceedings or review articles, leaving 36 original articles; 28 of these were in the English language. No additional articles were found from hand-searching uncited journals.

From the MEDLINE and BIDS-ISI search strategies combined, 54 articles were retrieved for EBCT. From these, 20 were not applicable to EBCT and 22 were abstracts from conference proceedings or review articles, leaving 12 original articles; 8 of these were in the English language. No additional articles were found from hand-searching uncited journals.

Of the articles located, the vast majority proved unsuitable for use in the review. In many cases they did not contain any reference to economics, except a brief mention in the abstract or conclusions in what were essentially clinical studies. Some articles discussed economic issues or made assertions about cost-effectiveness without presenting any economic data.

Only four of the 28 candidate articles in SCT could properly be described as economic studies, and only one of these approached acceptable standards for economic evaluation as judged by formal check-lists. For EBCT, only one true economic study was found and this was an analysis of costs.

Patient outcome and therapeutic impact

In these searches, the exclusion of review articles, abstracts and non-English language articles was performed for MEDLINE and BIDS-ISI, and publication dates were limited to pre-1997. This reduced the number of articles to be downloaded into the database for evaluation, but did not allow calculation of the numbers excluded at that stage. On combining those retrieved from MEDLINE and BIDS-ISI, 128 SCT and 165 EBCT articles were identified. After reading all these abstracts, 35 SCT and 18 EBCT were sufficiently applicable to the topic for retrieval of the full article. The reasons for excluding 93 SCT and 147 EBCT studies were because they did not meet either the preliminary criteria or one of the more specific exclusion criteria given in chapter 3.

On reading the full articles, 16 of the 35 SCT and seven of the 18 EBCT studies were excluded for one of the criteria already described, and 16 SCT and five EBCT studies were excluded for reasons shown in chapter 5. The remaining three SCT studies supplied information on therapeutic impact and, of the remaining six EBCT studies, three supplied information on therapeutic impact and three on patient outcome.

Diagnostic impact

No studies designed specifically to evaluate diagnostic impact were found during the search for therapeutic impact and patient outcome studies.

Twenty-one studies comparing the role of latest generation CT with alternative technologies were found (11 SCT, seven PE and three EBCT). Not all of these fulfilled the inclusion criteria. Of the 11 in the SCT liver lesion category, two duplicated studies were excluded. The nine remaining that concerned the identification of liver lesions used three comparative technologies: MRI techniques (three studies), US and MRI (two) and conventional CT (four). For PE, it was primarily the role of SCT with that of VP that was being compared (four studies). Only one of these compared the two modalities with an independent reference standard and was included in the review. Of the other three, one study compared echocardiography, one compared an MRI time-of-flight technique, and one compared the accuracy of MRI for detecting PE. However, these studies were excluded owing to insufficient data and the small number of patients (< 10) in the MRI study, the MRI time-of-flight study was an animal study, and the echocardiography study did not evaluate SCT with an independent reference test; it used SCT as the reference standard.

In EBCT, three comparative studies were identified, each comparing different technologies: intravascular US, ECG and thallium exercise test, and fluoroscopy. The intravascular US study performed an *in-vitro* analysis and was therefore excluded.

TABLE 10 *Number of articles retrieved from MEDLINE and BIDS-ISI, and the number of additional articles retrieved from the other resources*

Search resource	No. articles	
	SCT	EBCT
MEDLINE and BIDS-ISI	2166	1119
Additional articles		
Reference lists	0	1
EMBASE	0	0
Inside Information Plus	0	0
FirstSearch	0	0
System for Information on Grey Literature	0	0
Manufacturers	0	0
Any source	2166	1120

TABLE 11 *Number of articles excluded after application of preliminary inclusion criteria and number remaining*

Criterion	SCT	EBCT
No. articles excluded		
Published in 1997	296	33
Abstract	416	298
Review article	180	112
Non-English language	456	98
Case report	15	5
Editorial	11	3
Letter	27	12
Articles remaining	930	585

TABLE 12 *Number of articles excluded after use of major application-specific inclusion criteria*

Inclusion criteria	SCT – liver	SCT – PE	EBCT
Anatomical location	Liver	Lung	Heart
Type of disease	Cancer	Acute PE	CAD
Investigation	SCT phase or protocol	Diagnostic accuracy	Prediction, accuracy or reproducibility
No. articles excluded	881	914	530

Diagnostic performance

Electronic searches

Both MEDLINE and BIDS-ISI were searched from 1981 to the end of 1996. The combination of these two searches retrieved 2166 SCT and 1119 EBCT articles. The output from each of the additional searching methods was compared with those from MEDLINE and BIDS-ISI. One additional reference on EBCT was found (*Table 10*).

On applying the initial set of inclusion criteria shown in *Table 11*, 1236 SCT and 535 EBCT studies were excluded; 930 SCT and 585 EBCT studies remained. The classifications shown are not exclusive; for example, a review article can also be counted as a non-English language article.

When the subject-specific inclusion criteria of *Table 12* were applied to the information provided in the abstract, these criteria reduced the numbers that were of potential value from 930 to 49 for SCT for liver lesions and 16 for SCT for PE, and from 585 to 55 for EBCT studies.

Full copies of these articles were acquired and rechecked against all the criteria previously described, with further exclusions made as necessary. This process excluded a further 20 SCT liver lesion articles, nine for SCT for PE, and five for EBCT. The results of applying the subject-specific inclusion criteria are summarised in *Table 13*.

On the basis of these application-specific criteria, ten SCT liver lesion articles, three concerning SCT for PE and 24 on EBCT were excluded for the reasons shown in chapter 5, leaving the 19 SCT liver lesion studies, four SCT PE studies and 26 EBCT studies included in the review. The application of the final set of quantitative inclusion criteria identified seven symptomatic EBCT studies for inclusion in the quantitative analysis, excluding the other ten symptomatic EBCT studies for reasons identified in chapter 5.

Manufacturers

Replies were obtained from Siemens, Toshiba and Philips. Toshiba gave no information; Siemens supplied 59 references on SCT and ten on EBCT but no applicable references were previously unknown; Philips supplied 12 clinical publications on SCT, none of which was applicable to the review topic.

Grey literature

A few non-English language PhD theses and medical evaluation reports were retrieved from

TABLE 13 *Number of articles excluded after use of application-specific inclusion criteria*

Specific application	Criterion	No. studies excluded
All applications	Ten or fewer patients	4
	Non-human study	9
	Duplicate patient data set	7
HAP vs PVP	Information on both phases provided	0
HAP for HCC	Reference standard used	0
Protocol	Comparative study – one variable assessed	5
In place of VP	Reference standard used	0
With VP	Reference standard used	0
Prediction	Patient follow-up > 1 year	9
Diagnosis of CAD	Reference standard used and calcium score reported	3
Reproducibility	Time elapsed between scans < 1 week	1

the System for Information on Grey Literature, but none met the inclusion criteria.

Radiation dose

For SCT, of the 2166 articles within the database, 213 were retrieved by using the dose keyword search. After excluding review articles, abstracts etc., 103 studies remained. From reading the abstracts of these studies: 20 were not SCT; nine were retrieved when keywords were used in a different context than intended by our search; 17 were therapeutic doses; and one was a contrast medium dose. This left 56 articles for which the full text was required for further evaluation. Of these, a further 23 were excluded after rechecking against the same criteria, leaving 33 articles. After applying the criteria given in chapter 3, which required a study to include a comparison with another modality, eight remained and were included in the review of SCT radiation dose.

For EBCT, of the 1120 articles within the database, 102 were retrieved by using the dose keyword search (appendix 1); 74 remained after excluding review articles etc. On reading the abstracts: 54 were found not to concern EBCT; in five, keywords were used in a different context to that intended by our search; one was a review article; and nine were studies evaluating the dose of contrast media or therapeutic applications. This left five articles for which the full text was required for further evaluation. Three of these were included in the review.

Health economics

Articles covering any clinical application were eligible for inclusion in this category.

SCT

48. Kaneko M, Eguchi K, Ohmatsu H, Kakinuma R, Naruke T, Suemasu K, *et al.* Peripheral lung cancer: screening and detection with low-dose spiral CT vs radiography. *Radiology* 1996; **201**:798–802.

Twice-yearly examinations by both methods were carried out over a 2-year period on 1369 individuals who were at high risk of lung cancer. More cancers were detected by SCT than by radiography, and the overall detection rate was higher than in the previous period when SCT was not available. A rough costing analysis was carried out using payments by individuals to join the screening programme. The average costs per tumour detected were lower after the introduction of SCT. Marginal cost-effectiveness ratios were not calculated. The limitations of this preliminary analysis were recognised. This article indicated that data may be available for a more comprehensive analysis of the impact on treatment and outcome.

49. Lindgren BW, Demos T, Marson R,Posniak H, Kostro B, Calvert D, *et al.* Renal computed tomography with 3-dimensional angiography and simultaneous measurement of plasma contrast clearance reduce the invasiveness and cost of evaluating living renal donor candidates. *Transplantation* 1996;**61**:219–23.

Twenty-three living renal donor candidates were evaluated by using CT/three-dimensional CTA/plasma contrast clearance and intravenous urography/renal arteriography/creatinine clearance. The former strategy identified more relevant pathologies with less discomfort and inconvenience to the potential donors. The costs of the diagnostic strategies were compared by using charge data from the investigators' own hospital. The approach using three-dimensional CTA is less costly and, although no formal

measurement of effectiveness was carried out, it is presumed to be the dominant strategy.

50. Semelka R, Schlund J, Molina P, Willms AB, Kahlenberg M, Mauro MA, *et al.* Malignant liver lesions: comparison of spiral CT arterial portography and MR imaging for diagnostic accuracy, cost, and effect on patient management. *J Magn Reson Imaging* 1996;**1**:39–43.

CT during arterial portography (CTAP) and MRI were carried out within 1 week in each of 26 patients. Surgical data were evaluated for ten patients. CTAP was found to be more sensitive but less specific than MRI, and, in the judgement of the referring surgeon, CTAP did not alter management decisions based on the MRI findings. The costs for each procedure were estimated from billing data at the study hospitals and included overnight inpatient stays. On the basis of this limited information it is concluded that CTAP is more costly than MRI and does not add to the accuracy of diagnosis or the effectiveness of treatment.

19. van Erkel A, van Rossum A, Bloem JL, Kievit J, Pattynama PM. Spiral CT angiography for suspected pulmonary embolism: a cost-effectiveness analysis. *Radiology* 1996;**201**:29–36.

This study used a decision analytical model to compare SCT with conventional CTA using the diagnostic accuracy and prognosis drawn from the literature. The costs of investigations and therapies were taken from the investigators' hospital. Multiple diagnostic strategies were compared by using lives saved and cost per life saved as outcome measures. The study was well conducted and clearly presented. The need to use life years saved and some adjustment for QoL in the outcome measure was recognised, but data were not readily available.

EBCT

51. Hernigou A, Perrin JP, Grataloup C,Philippe E, Plainfosse MC. Cost comparison of electron beam tomography with conventional computed tomography scanning. *Acad Radiol* 1996;**3**:5145–6.

The costs of acquisition, installation, maintenance and operation of an EBCT unit were compared with those of CT units with similar levels of activity. Capital costs were depreciated over 7 years. Comparative cost figures were not clearly presented, with the average cost per image acquired (Fr.fr.10.36 for EBCT, Fr.fr.7.38 for conventional CT) being quoted without supporting details. This short note provided some basic information on EBCT costs, but it did not constitute a proper comparative study.

Patient outcome

Articles covering any clinical application were eligible for inclusion in this category.

SCT

No studies were eligible for inclusion in the review.

EBCT

52. Stanford W, Travis ME, Thompson BH, Reiners TJ, Hasson RR, Winniford MD. Electron beam computed tomographic detection of coronary calcification in patients undergoing percutaneous transluminal coronary angioplasty: predictability of restenosis. A preliminary report. *Am J Card Imaging* 1995;**9**:257–60.

Twenty patients were studied to evaluate whether the presence of CAC detected by EBCT was predictive of re-stenosis after percutaneous transluminal coronary angioplasty (PTCA). EBCT was performed immediately before, immediately after, and at 2–18 months after PTCA. Matching the arterial site for the measurement of coronary calcification was achieved by estimating the distance between the stenosis and the ostium of the coronary vessel affected. Restenosis was defined as a recurrence of symptoms and a reduced arterial diameter of 50% or more. The results showed that CAC at the PTCA site was significantly greater in re-stenosed than in non-stenosed patients. The authors indicated that EBCT may therefore predict outcome after PTCA but that a larger population study is needed.

53. Garden AS, Morrison WH, Clayman GL, Ang KK, Peters LJ. Early squamous cell carcinoma of the hypopharynx: outcomes of treatment with radiation alone to the primary disease. *Head Neck* 1996;**18**:317–22.

This retrospective study analysed the outcome of treatment with radiotherapy in 82 patients with early stage T1/T2 hypopharyngeal squamous cell carcinoma. EBCT was performed as part of the staging work-up in 36 patients. The influence of EBCT in patients with T2 disease was evaluated. Patients who had undergone EBCT had a 2-year actuarial local control rate of 83% compared with 71% ($p = 0.1$) in the group without CT. EBCT was one of a number of factors that may have contributeed to the improvement seen in the results in the latter half of the study.

54. Wong ND, Detrano RC, Diamond G, Rezayat C, Mahmoudi R, Chong EC, *et al.* Does coronary artery screening by electron beam computed tomography motivate potentially beneficial lifestyle behaviors? *Am J Cardiol* 1996;**78**:1220–3.

This behavioural study evaluated the impact of knowledge of CAD newly diagnosed by EBCT on cardiovascular risk-reducing behaviours. A total of 703 asymptomatic individuals (560 male and 143 female) from a larger study population of self-referrals who underwent EBCT, and for whom baseline risk factors and follow-up questionnaires were available, formed the study population. Health behaviour factors at follow-up were significantly associated with higher calcium scores. The authors conclude that when baseline risk factors associated with high calcium scores are corrected for, a number of risk-reducing behaviours are associated with increased calcium scores on EBCT.

Therapeutic impact

Articles covering any clinical application were eligible for inclusion in this category.

SCT

55. Lacrosse M, Trigaux JP, van Beers BE, Weynants P. 3D spiral CT of the tracheobronchial tree. *J Comput Assist Tomogr* 1995;**19**:341–7.

Patients who had undergone conventional CT regarding tracheobronchial status went on to high-resolution SCT imaging and three-dimensional rendering of the CT data. Two-dimensional SCT images were read by consensus of two experienced thoracic radiologists. Three-dimensional processing and multiplanar reformats were then produced and read. The relevance of any additional information to patient management of the three-dimensional processing was established by consensus comparison with findings at endoscopy and a second retrospective reading of the two-dimensional slices. No statistical analysis was possible because of the small heterogeneous study population and bias in the study. The authors described three of 11 instances in which diagnostic information from SCT led to major modifications in patient management.

50. Semelka R, Schlund J, Molina P, Willms AB, Kahlenberg M, Mauro MA, *et al.* Malignant liver lesions: comparison of spiral CT arterial portography and MR imaging for diagnostic accuracy, cost, and effect on patient management. *J Magn Reson Imaging* 1996;**1**:39–43.

This study compared spiral CTAP and MRI for diagnostic accuracy, procedural cost, and effect on patient management. Twenty-six consecutive patients with suspected limited liver disease determined by dynamic or spiral contrast enhanced CT, who were candidates for liver resection, underwent CTAP and MRI. Images were interpreted prospectively in a blinded study design by experienced investigators to determine the presence of liver lesions and segmental involvement.

The results showed that the CTAP findings did not change patient management in any of the cases over MRI, but MRI did change patient management in seven cases over CTAP ($p < 0.015$).

56. Winston CB, Wechsler RJ, Salazar AM, Kurtz AB, Spirn PW. Incidental pulmonary emboli detected at helical CT: effect on patient care. *Radiology* 1996;**201**:23–7.

This was a retrospective study of changes in patient management resulting from incidental findings at SCT of the chest. A computer search of 1879 consecutive contrast enhanced thoracic SCT studies identified 28 patients in whom the diagnosis of PE was made or suspected. These 28 CT studies were reviewed by three radiologists. In 18 patients in whom PE was not suspected, intraluminal defects were confirmed. PE was an incidental finding in 1% of these patients. Clinical management was changed in 11 patients, who received anticoagulants or caval filter placement as a result of the CT findings.

EBCT

57. Barloon TJ, Galvin JR, Mori M, Stanford W, Gingrich RD. High-resolution ultrafast chest CT in the clinical management of febrile bone marrow transplant patients with normal or nonspecific chest roentgenograms. *Chest* 1991;**99**:928–33.

Thirty-three bone marrow transplant recipients with suspected pulmonary disease underwent EBCT when chest radiography failed to provide sufficient information to initiate or continue treatment. The referring physician completed a clinical management data sheet after notification of the radiological result. EBCT changed management in three out of 14 patients with normal chest radiographs, yielding no significant benefit ($p < 0.05$). EBCT changed management in nine of 22 patients with non-specific chest radiographs and, in a further eight patients, it provided additional information that did not change management, yielding a significant benefit over conventional chest radiography ($p < 0.001$). Patient outcome was tabulated for death/recovery but this information was not linked to the CT findings.

58. Mousseaux E, Hernigou A, Azencot M, Sapoval M, Auguste M, Gaux JC. Evaluation by electron beam computed tomography of intracardiac masses suspected by transoesophageal echocardiography. *Heart* 1996;**76**:256–63.

Consecutive patients with suspected or known intracardiac masses assessed by transoesophageal US were examined by EBCT. Those with echo features of thrombus were excluded. EBCT showed no cardiac mass in 20 patients in whom anatomical structures were seen that explained the US findings. EBCT detected pathological lesions in 56 of 76 patients. The study analysed the impact of additional information from EBCT over US on decision making in terms of confirming the presence of a mass or demonstrating the extent and anatomy of a tumour. The influence of EBCT on patient management was assessed by reviewing the patients' notes with the referring clinician/echocardiographer. In 53 patients, EBCT modified or confirmed the diagnosis through lesion characterisation or anatomical information. EBCT contributed to surgical planning in 17 but not in nine, and to the decision against surgery in 55. There were no data on health outcomes.

59. Szolar DH, Groell R, Braun H, Preidler K, Stiskal M, Kern R, *et al.* Ultrafast computed tomography and three-dimensional image processing of CT sialography in patients with parotid masses poorly defined by magnetic resonance imaging. *Acta Otolaryngol* 1996;**116**:112–18.

Patients with poor-quality MRI studies were referred for EBCT evaluation. Three-dimensional surface reconstruction of CT data was also performed in nine patients. Images were interpreted by two experienced radiologists. Without knowledge of the surgical/pathological results, both imaging modalities were compared, and imaging was compared with clinical results and post-operative findings. MRI and EBCT were of equal value in defining tumour location. EBCT better defined tumour margins. A scoring system was used to compare anatomical detail and the display of ductal pathology. Three-dimensional reconstruction of EBCT data provided an improved display of ductal pathology and anatomical detail compared with two-dimensional CT images ($p < 0.05$). In two out of 13 patients, surgical management was changed by three-dimensional reconstructed EBCT images.

Diagnostic impact

Articles covering any clinical application were eligible for inclusion in this category.

SCT – liver lesions

No studies designed specifically to address the diagnostic impact of SCT were identified for any clinical application. However, a few compared the role of SCT to that of alternative imaging modalities in studies designed to assess the diagnostic performance in our selected clinical applications.

For SCT for liver lesions, although studies shared similar flaws in methodology such as the use of several inadequate reference standards, they did supply some indirect evidence of the comparative accuracy of the modalities because comparisons were based on the same standards. These studies are described next.

SCT liver lesion studies designed to compare modalities

Studies comparing SCT with alternative modalities

60. Jung G, Krahe T, Krug B, Hahn U, Raab M. Delineation of segmental liver anatomy. Comparison of ultrasonography, spiral CT, and MR imaging for preoperative localization of focal liver lesions to specific hepatic segments. *Acta Radiol* 1996;**37**:691–5.

Twenty-four patients with a range of tumour types were identified (mainly metastases from colon carcinoma) by intraoperative US and palpation. SCT was performed using a 5 mm slice, an 8 mm/s table speed, and a 5 mm increment, after the administration of 100 ml of contrast medium at 2 ml/s with a 65 s delay. Twenty-two patients also underwent MRI with a 1.5T system: 8 mm slice, 10 mm intervals, T1-weighted (TR/TE 450/20 ms) and T2-weighted (TR/TE 1800/80/20 ms) spin echo technique. All patients underwent an abdominal US scan using a high-resolution real-time scanner with a 3.5 MHz transducer. SCT provided the best localisation of lesions to specific segments and correctly described the primary segmental location in 21 of 24 patients, followed by MRI with 17 of 22 and US with 15 of 24. All examinations were performed within 11 days of surgery. CT and MRI were interpreted by two radiologists blinded to the US and operative results.

61. Oi H, Murakami T, Kim T, Matsushita M, Kishimoto H, Nakamura H. Dynamic MR imaging and early-phase helical CT for detecting small intrahepatic metastases of hepatocellular carcinoma. *AJR Am J Roentgenol* 1996;**166**:369–74.

This study compared the performance of dynamic intravenous Gd-DTPA-enhanced MRI, spin echo MRI, HAP CT and delayed phase (DP)

CT. Iodised oil CT after transcatheter arterial chemoembolization was used as a reference standard. Forty-six patients had 225 intrahepatic metastases of HCC less than 3 cm in diameter. Most lesions were detected with dynamic MRI (62%), with the performance of HAP SCT coming second (47%). Dynamic MRI was significantly ($p < 0.001$) better than HAP SCT only for tumours of < 10 mm. Combining the two MRI and the two CT techniques, MRI was significantly ($p < 0.001$) better than CT for tumours of less < 10 mm (67% versus 52%).

62. Ueda K, Kitagawa K, Kadoya M, Matsui O, Takashima T, Yamahana T. Detection of hypervascular hepatocellular carcinoma by using spiral volumetric CT: comparison of US and MR imaging. *Abdom Imaging* 1995;**20**:547–53.

This study compared precontrast SCT, arterial phase SCT, late phase SCT, US imaging and unenhanced spin echo MRI in the detection of HCC. Angiography including SCT during arteriography and arterial portography was used as a reference standard. Forty-three HCCs were identified. HAP SCT identified five lesions alone; MRI identified three lesions alone. There was no statistical difference between the performance of MRI and CT or MRI and US. All phases of CT together performed statistically better than US ($p < 0.05$).

63. van Hoe L, Bosmans H, Aerts P, Baert AL, Fevery J, Kiefer B, *et al.* Focal liver lesions: fast T2-weighted MR imaging with half-Fourier rapid acquisition with relaxation enhancement. *Radiology* 1996;**201**:817–23.

Half-Fourier rapid acquisition with relaxation enhancement (RARE) was compared with multishot RARE and SCT for motion artefact, anatomical sharpness, overall image quality and lesion conspicuity. Uniphasic CT at the peak hepatic enhancement was used for hypovascular lesions and biphasic CT (25 s and 80 s delays) was used for uncertain or hypervascular lesions. The results were combined. (CT slice thickness 5 mm, pitch 1.0 or 1.5, 120 kV, 292 mA, 0.75 s, 125 ml of 350 mgI/ml at 2.5 ml/s.) CT performed better than both half-Fourier and multishot RARE in terms of motion artefact, anatomical sharpness and overall image quality. Half-Fourier RARE had better conspicuity than SCT for haemangiomas (five of 21 better), SCT had better conspicuity for solid lesions (11 of 54 better), and both techniques were equivalent for the conspicuity of cysts (23 equal). Differences between the numbers of lesions were observed only for metastases: CT depicted 49 of 50; half-Fourier RARE depicted 47 of 50; and multishot RARE depicted 38 of 50.

64. Yamashita Y, Mitsuzaki K, Yi T, Ogata I, Nishiharu T, Urata J, *et al.* Small hepatocellular carcinoma in patients with chronic liver damage: prospective comparison of detection with dynamic MR imaging and helical CT of the whole liver. *Radiology* 1996;**200**:79–84.

Triphasic SCT was performed with a 7 mm or 10 mm slice, pitch 1 or 1.5, 120 kV and up to 240 mAs. Delay times were 25 s arterial, 60–65 s portal, and 180 s DP. A dose of 2 ml/kg of iopamidol 300 was administered at 4 ml/s. For comparison, dynamic gadolinium-enhanced MRI was performed on a 1.5T system with a T1-weighted, fast low-angle shot sequence (150/4.8, flip 75), 8–10 mm slice, 0.8–1.0 mm gap, 128 × 256 matrix, and 26 × 34 cm field of view. Magnetic resonance images were separated into the three hepatic phases. A total of 72 HCCs in 27 patients were identified from MRI and CT images. ROC analysis was performed with three readers. Diagnostic ability was significantly ($p < 0.05$) better for arterial phase MRI than arterial phase CT for all three readers. DP SCT was significantly ($p < 0.05$) better than DP MRI. No difference was found between portal phase images. The difference in performance between MRI and CT was most significant for lesions of less than 2 cm.

Studies comparing SCT with conventional CT

65. Bonaldi VM, Bret PM, Reinhold C, Atri M. Comparison of helical and conventional computed tomography of the liver. *Can Assoc Radiol J* 1995;**46**:443–8.

One hundred and sixty-six patients were assigned randomly to SCT (79 patients) or conventional CT (87 patients) to compare enhancement and image quality rather than lesion detection performance. CT parameters were optimised for habitus: 140 or 120 kV and 150 or 300 mAs for SCT; 120 kV and 250, 300, 400 or 500 mAs for conventional CT. Both techniques were performed with 10 mm slices. Subgroups of patients were clinically assigned to one of three groups of non-ionic or one ionic contrast medium. A volume of 2 ml/kg was administered at a rate of 3 ml/s for SCT and using a biphasic injection with half at 2.0 ml/s and half at 0.7 ml/s for conventional CT. Scanning was initiated 58–66 s for SCT or 45–65 s for conventional CT after the beginning of the injection. Liver enhancement was significantly ($p < 0.05$) greater with SCT for the ionic contrast medium

and two of the three non-ionic contrast media used. Image quality was not significantly different.

66. Choi BI, Shin YM, Han JK, Chung JW, Park JH, Han MC. Focal hepatic nodules after transcatheter oily chemoembolization; detection with spiral CT versus conventional CT. *Abdom Imaging* 1996;**21**:33–6.

Forty-two patients underwent transcatheter oily chemoembolization for suspected HCC. Conventional CT parameters were: 10 mm slice; 10 mm intervals; 1 s scan; 120 kV; 220 mA. The SCT parameters were: 10 mm slice; pitch 1; 24 s acquisition time; 120 kVp; 210 mAs. A total of 107 nodules were identified by two radiologists with SCT compared with 98 with conventional CT. In six patients, more lesions were identified with SCT; of these, two had no nodules on conventional CT. Nine more lesions were detected with SCT than with conventional CT ($p = 0.002$). All nine were less than 20 mm in diameter; seven were less than 10 mm.

67. Fujita M, Kuroda C, Kumatani T, Yoshioka H, Kuriyama K, Inoue E, *et al.* Comparison between conventional and spiral CT in patients with hypervascular hepatocellular carcinoma. *Eur J Radiol* 1994;**18**:134–6.

This study compared conventional CT and arterial phase SCT for the detection of HCC. Hepatic digital subtraction angiography was used as a reference standard, indicating 56 HCCs (size 6–90 mm, mean 25 mm) in 29 patients. Conventional CT was performed with the following parameters: 10 mm slice; 120 kV; 1 or 2 s scan; 250 or 200 mA; 100 ml of 300 mgI; biphasic injection 50 ml at 0.7 ml/s and 50 ml at 0.3 ml/s; 70 s scan delay from beginning of injection. SCT parameters were: 10 mm slice; pitch 1; 120 kV; 210 mA; 1 s scan; 100 ml at 2 or 3 ml/s; 40–45 s scan delay from beginning of injection. Neither technique detected any of the seven tumours of < 10 mm. Conventional CT did not detect any tumours that SCT did not, whereas SCT detected 12 tumours that conventional CT did not. A significant difference between conventional and SCT for HCC detection was found only for tumours of between 10 and 20 mm in size ($p = 0.01$).

68. Polger M, Seltzer SE, Head BL, Savci G, Silverman SG, Adams DF. Spiral computed tomography of the liver – contrast agent pharmacokinetics and the potential for improved hepatic enhancement. *Acad Radiol* 1995;**2**:19–25.

In this randomised controlled trial, patients were randomised to two control groups receiving conventional CT and five experimental groups receiving SCT. One of the experimental groups had identical parameters, enabling direct comparison with the conventional CT control group. These parameters were a monophasic dose of 150 ml diatrizoate meglumine at a rate of 2.5 ml/s with a 60 s delay. Liver enhancement in these two groups was significantly less ($p = 0.011$) for conventional CT compared with SCT in the last slice in the data set. No difference was found between other experimental groups when using different delay times and doses, or between these SCT groups and the conventional CT control groups.

SCT – PE

Seven studies were identified that evaluated the performance of other modalities and that of SCT, with both compared with the same reference standard. Six of these were excluded for the reasons shown chapter 5. In the remaining study, a patient subgroup received both VP scanning and SCT and the comparison was made with the independent reference of angiography.

69. Remy-Jardin M, Remy J, Deschildre F, Artaud D, Beregi JP, Hossein-Foucher C, *et al.* Diagnosis of pulmonary embolism with spiral CT: comparison with pulmonary angiography and scintigraphy. *Radiology* 1996;**200**:699–706.

A subset of 25 of 75 patients underwent VP scanning within 3 days of CT or angiography. Of 13 patients shown to be negative for PE by angiography, VP classified two as normal, six as low probability, three as intermediate and two as high. SCT identified all except one as negative, with the incorrect diagnosis being from an inconclusive scan. Of 12 patients who were positive with angiography, CT classified all of them correctly. VP classified one as normal, three as low probability, four as intermediate and four as high probability for PE.

EBCT – CAD

No studies designed specifically to address the diagnostic impact of EBCT were identified.

For the specific application of the comparative accuracy of diagnosing CAD, three studies were identified that compared three modalities. One *in-vitro* study was excluded. The remaining two studies are summarised next.

22. Agatston AS, Janowitz WR, Hildner FJ, Zusmer NR, Viamonte M, Detrano R. Quantification of coronary artery calcium using ultrafast computed tomography. *J Am Coll Cardiol* 1990;**15**:827–32.

A subgroup of 50 patients underwent both fluoroscopy and EBCT with comparison with angiography. On angiography, 18 had normal or non-obstructive disease and 32 had obstructive CAD. Forty-five of 50 patients had calcium reported by EBCT compared with 26 of 50 by fluoroscopy. Using a calcium score greater than zero, EBCT was 100% sensitive but with a low specificity of 28%, whereas fluoroscopy had a better balance with 75% sensitivity and 56% specificity.

70. Kajinami K, Seki H, Takekoshi N, Mabuchi H. Noninvasive prediction of coronary atherosclerosis by quantification of coronary artery calcification using electron beam computed tomography – comparison with electrocardiographic and thallium exercise stress test results. *J Am Coll Cardiol* 1995;**26**:1209–21.

A comparison with angiography of EBCT, ECG and thallium exercise stress test results for 251 patients was performed. The results were presented for all patients as well as being subclassified by age in decades for each modality and by log calcium score cut-off for EBCT. The complementary and combined performance of ECG and thallium tests was also presented. Sensitivity and specificity values for each modality, including the two complementary and combined tests, were compared for any significant difference by using the McNemar test. The sensitivity of EBCT for all subgroups by age as well as the total group was not significantly different ($p > 0.2$) from that of ECG. The specificity of EBCT was higher than that of ECG for patients aged 40–50 years ($p = 0.076$), 60–70 years ($p = 0.015$) and all patients ($p = 0.058$). The sensitivity of EBCT was lower than that of the thallium test for patients aged 60–70 years ($p = 0.17$). The specificity of EBCT was higher than the thallium test for all groups except patients aged > 70 years, in whom no difference was found. The best results of the ECG and thallium tests were achieved by using the results together (i.e. positive for at least one of the tests). The comparison of both these tests with EBCT illustrated that EBCT had higher specificity in all groups except patients > 70 years of age, but EBCT had lower sensitivity in patients aged 60–70 years ($p = 0.0035$) and also in the total patient group ($p = 0.001$).

Diagnostic performance

Only articles on our chosen clinical applications were eligible for inclusion in this category.

SCT – liver lesions

No previous systematic reviews were found.

No attempt will be made to combine the results from these articles for a number of reasons. We initially sought to include studies that used the single gold standard of resection and pathology on every patient, but none was found. Pathological confirmation on selected lesions, often by needle biopsy rather than on a resected specimen, is more commonly used. In this situation it is not possible to quote true sensitivity and specificity values because not every lesion seen on a scan is biopsied. Instead, all those with similar imaging appearances will be assumed to have the same pathological results. The main feature that the studies have in common is the lack of a single gold standard; the confirmation of tumour presence and type is reported using a range of other investigations.

These studies fall into the following three categories:

- studies comparing HAP with PVP for the detection of any type of lesion
- studies using HAP to visualise hypervascular HCC compared with a reference standard
- studies designed to compare protocols.

Studies comparing HAP with PVP for lesion detection

Four studies were identified in this category.[71–74] A wide range of tumour types were represented in this group, including HCC in a study by Oliver *et al.*,[71] which also provided information for the second category. Delay times and other protocol parameters varied between studies, making direct comparisons difficult. Details are provided in *Table 14*.

Studies using HAP to visualise hypervascular HCC tumours with comparison to a reference standard

Four studies, including the study by Oliver *et al.*[71] (described above), were identified.[61,62,71,75] A single reference standard was not used in any study; instead, a variety of techniques provided verification. In addition, differences in HAP delay times and scanning protocols make direct comparison difficult. These study details are supplied in *Table 15*, except for the study by Oliver *et al.*[71] (see *Table 14*). No information on tumour size was provided, so there is no additional information to that already presented.

Studies designed to compare protocols

No studies were found that concentrated on comparing protocols for HAP and/or PVP for their performance in detecting liver lesions. There were, however, studies that focused on the PVP,

TABLE 14 *Details of studies using HAP to visualise HCCs and comparing with PVP*

	Bonaldi *et al.* 1995[72]	**Hollett *et al.* 1995[73]**	**Mitsuzaki *et al.* 1996[74]**	**Oliver *et al.* 1996[71]**
Centre	Quebec, Canada	Stanford, CA	Kumamoto, Japan	Pittsburgh, PA
No. patients	103	27 (total 96)	67 (total 227)	81 (42 proven)
Male (%)	57	64 (total 96)	74 (total 227)	68
Age (years)	20–88 (av. 54)	18–85 (av. 60) (total 96)	32–38 (av. 63) (total 227)	37–73 (av. 52)
Weight (kg)	36–102 (av. 65)	N/S	N/S	N/S
Symptoms	N/S	N/S	Liver cirrhosis	Advanced cirrhosis
Tumour types	Large range	Variety	Hepatomas	HCCs
Tumour size (cm)	0.5–2.5 (av. 1.7) (HAP-only lesions)	≤ 1.5	av. 1.9 (2 ml/s) av. 2.0 (4 ml/s)	N/S
Tumour cut-off	None	20	5	10
No. lesions per patient	1.2	N/S	1.6	3.7 (42 proven)
Unenhanced scan	10 mm slice 20 mm interval	10 mm slice 20 mm interval	N/P	5 mm slice 8 mm interval
HAP delay (s)	15	20	1) 30 (at 2 ml/s) 2) 35 (at 2 ml/s) 3) 20 (at 4 ml/s) 4) 25 (at 4 ml/s)	28 (at 2.5–3.0 ml/s) 20 (at 5.0 ml/s)
PVP delay (s)	90	51–72 60 ± 7	80 (at 2 ml/s) 60 (at 4 ml/s)	70 60
Late delay (s)	N/P	N/P	240 (at 2 ml/s) 180 (at 4 ml/s)	N/P
Slice (mm)	10	7	7–10	7
Pitch	1.0, 0.5	1.0, 1.3, 1.6	1.0	1.0
Contrast dose	2 ml/kg	150 ml	2 ml/kg	150 ml
Injection rate (ml/s)	6	5	2 4	2.5–3.0 ionic 2.5–5.0 non-ionic
Contrast medium	Meglumine iothalamate 60 Ioversol 320	Iohexol 300	Iopamidol 300	Meglumine 60 iothalamate 60 Ioversol 320
Scanner	TCT 900SX	HiSpeed Advantage	HiSpeed Advantage	HiSpeed Advantage
Reconstruction (mm)	10	7–10	7	N/S
Interpolation	360°	N/S	N/S	N/S
Voltage and current	120–140 kV 150–300 mAs	120 kVp, 251 ± 25 or 233 ± 27 mAs	120 kVp	N/S

N/S, not stated; N/P, not performed

continued

TABLE 14 contd *Details of studies using HAP to visualise HCCs and comparing with PVP*

	Bonaldi *et al.* 1995[72]	**Hollett *et al.* 1995[73]**	**Mitsuzaki *et al.* 1996[74]**	**Oliver *et al.* 1996[71]**
Viewing and windowing	FOV 320 mm (71 patients) 400 mm (32 patients)	(400/40) and (150/50–80)	Photographed (150/50)	N/S
No. observers	2	3	N/S	3
Blinded	Yes	No	No	Yes
Reference standards: number (%)	None	Histology: 13 Biopsy: 12 Metastases: 8 Follow CT: 5	Histology: 23 Biopsy: 42 Follow CT: 44	Histology: 16 (20) Biopsy: 65 (80) Follow CT: 26 (40) Proven: 42 (52)
Conclusions: number (%)	***119 lesions*** HAP only: 9 (8) PVP only: 40 (34) Both: 70 (59) ***103 patients*** HAP only: 5 (5) PVP only: 21 (20)	***? Lesions*** HAP only: 33 ***27 patients*** HAP only: 6 (22) ***Exclusive patients*** HAP only: 1 (4)	***All 4 protocols:*** ***109 lesions*** HAP only: 9 (8) PVP only: 1 (1) DP only: 1 (1) HAP: 104 (95) PVP: 64 (59) DP: 91 (83)	***157 proven lesions*** HAP only: 18 (11) Un only: 3 (2) Both: 9 (6) ***42 proven patients*** HAP only: 10 (24) Un only: 3 (7) Both: 7 (17)

N/S, not stated; Un, unenhanced

where the maximal parenchymal enhancement (which is generally believed to maximise conspicuity of hypovascular lesions) was sought, rather than a measure of detection performance. Dodd and Baron[76] reviewed the intrinsic (e.g. patient's weight) and extrinsic (e.g. contrast administration protocol) variables that may affect hepatic perfusion. They pointed out that studies that fail to control these factors are impossible to compare with one another.

A class of protocol investigations is the set of studies using automated scan technology to optimise the timing of the SCT protocol. Four studies by two author groups were identified that fitted into this category.

The details of all the studies included are described in *Tables 16–21*, which cover the study characteristics, the parameters investigated and the study conclusions, for both PVP protocols and automated scan technology.[68,77–86]

SCT – PE

No previous systematic reviews were found.

There are three proposed roles for SCT in the diagnosis of PE:

- together with other modalities as a primary diagnostic test
- in place of VP, as a primary diagnostic test
- after VP.

For this review, evidence was sought on the diagnostic accuracy of SCT in all three of these roles. To illustrate diagnostic accuracy, comparison with a gold standard is necessary, pulmonary angiography in this case. Study designs should follow the broad examples shown in *Figures 5* and *6* for each of the second and third roles respectively.

Four studies comparing the diagnostic performance of SCT with the gold standard of pulmonary angiography were included. The only other modality involved was VP scanning. All four studies addressed the role of SCT in patients with a clinical suspicion of PE. One study[20] did not involve VP scanning and used clinical symptoms to select the study group (*Figure 5*). The second[69] also followed the design of *Figure 5,* although a subset of the patient group also received VP scanning (this arm of the study was described earlier under diagnostic impact). Two studies[87,88] were primarily concerned with the central option displayed in *Figure 6,* indicating the role of SCT after indeterminate or inconsistent VP. Study details are given in *Table 22.*

In 1992, Remy-Jardin *et al.*[20] investigated 42 consecutive patients with both SCT and pulmonary angiography; all had clinically suspected pulmonary thromboembolism. For the central vessels only, the sensitivity and specificity for contrast-enhanced SCT were 100% and 96% respectively, values that were enhanced by the

TABLE 15 *Details of liver lesion studies with comparison to reference standard*

	Choi *et al.* 1996[75]	**Oi *et al.* 1996[61]**	**Ueda *et al.* 1995[62]**
Centre	Seoul, Korea	Osaka, Japan	Kanazawa, Japan
No. patients	58	46 (of 49)	43 (of 512)
Male (%)	88	80 (of 49)	75 (of 512)
Age (years)	39–77 (av. 54)	36–77 (av. 62 of 49)	29–84 (av. 55 of 512)
Weight (kg)	N/S	N/S	40–75 (av. 62)
Symptoms	47 Hepatitis B 42 Cirrhosis	49 Cirrhosis or hepatitis	251 Hepatitis B/C 186 Liver cirrhosis 9 Biliary cirrhosis 66 Alcoholic fibrosis
Tumour size (no.)	< 1 cm (31) 1–2 cm (23) 2–3 cm (21) > 3 cm (36)	< 1 cm (143) 1–2 cm (54) 2–3 cm (28)	5–9 mm (7) 10–14 mm (14) 15–19 mm (8) 20–29 mm (8) > 30 mm (6)
Lesions per patient	1–6 (mean 1.9)	mean 4.9	N/S
Slice thickness (mm)	10	8	8
Table speed (mm/s)	10	7	8
No. slices	20	23	24
Reconstruction (mm)	N/S	N/S	2
Scanner	Somatom Plus S	Somatom Plus	Somatom Plus
Voltage, current	120 kVp, 210 mA	N/S	120 kVp, 165 mA
Windowing	N/S	N/S	Cine 150 HU window
HAP delay (s)	35	38 (80 ml); 33 (90 ml)	30
Contrast protocol	120 ml at 3 ml/s	80 ml at 2 ml/s; 90 ml at 3 ml/s	70 ml at 2 ml/s
Contrast media	Meglumine ioglicate 300	Iohexol/iopamidol	300 mg/ml I
No. observers	2	N/S	2 or 3
Blinded	Yes	N/S	Yes
Reference standards	Surgery 21 Biopsy/iodised CT 37	Histology 11 Iodised oil CT 35	Histology 7 1 Year follow-up 36
Comparisons	DP (180 s)	MRI, DP (5 min)	MRI, US, DP (7 min)
Conclusions: number (%)	***HAP + DP*** < 1 cm: 22 (71) 1–2 cm: 20 (87) 2–3 cm: 20 (95) > 3 cm: 36 (100) Total: 98 (88) ***HAP*** < 1 cm: 20 (65) 1–2 cm: 20 (87) 2–3 cm: 18 (86) > 3 cm: 35 (97)	***HAP + DP*** < 1 cm: 52 (36) 1–2 cm: 41 (76) 2–3 cm: 25 (89) Total: 118 (52) ***HAP*** < 1 cm: 49 (34) 1–2 cm: 35 (65) 2–3 cm: 22 (82) Total: 106 (47)	***HAP + DP*** 5–9 mm: 4 (57) 10–14 mm: 12 (86) 15–19 mm: 8 (100) 20–29 mm: 8 (100) > 30 mm: 6 (100) Total: 38 (88) ***HAP*** 5–9 mm: 1 (14) 10–14 mm: 3 (21) 15–19 mm: 1 (13)

TABLE 16 *Studies comparing PVP protocols – study characteristics*

	Brink *et al.* 1995[77]	**Herts *et al.* 1995[78]**	**Hoeffel *et al.* 1996[79]**	**Kopka *et al.* 1995[80]**	**Polger *et al.* 1995[68]**	**Silverman *et al.* 1995[81]**	**Small *et al.* 1994[82]**
Centre	New York, NY	Cleveland, OH	Paris, France	Goettingen, Germany	Boston, MA	Washington, DC	Atlanta, GA
No. patients	487	169	121	75	131	25	20
Male (%)	68	43	66	45	47	40	100
Age (years)	19–83 (av. 61)	19–85 (av. 57)	20–83 (av. 53)	Per group only	23–83 (av. 55)	34–79 (av. 56)	19–50 (av. 29)
Weight (kg)	34–141 (av. 83)	N/S	45–92 (av. 64)	N/S	N/S	N/S	55–114 (av. 79)
Clinical condition	N/S	N/S	N/S	Metastases	N/S	Metastases	Normal
Scanner	Somatom Plus or Plus S	Somatom Plus or Plus S	Elscinct CT Twin	HiSpeed Advantage	Somatom Plus or Plus S	HiSpeed Advantage	Somatom Plus
Slice (mm)	8	10	6.5	10	10	10	5
Pitch	1	1	2:1	1	1	1	1
Reconstruction interval (mm)	10	10	5	N/S	10	N/S	N/S
Voltage (kV)	120	120	N/S	120	125	120	120
Current (mA)	165	165/210	N/S	300	165	300–320	210

exclusion of inconclusive CT findings. The authors suggested that SCT represents a rapid way to detect acute thromboembolic disease, especially in those at risk from the complications of pulmonary angiography. They pointed out that SCT may fail to depict small peripheral emboli.

Remy-Jardin *et al.*,[69] in 1996 investigated a group of 75 patients with both SCT and pulmonary angiography, 25 of whom also underwent VP scanning. All were under suspicion for acute PE. For the central vessels only, the sensitivity and specificity for contrast-enhanced SCT were 91% and 78% respectively. They evaluated subsegmental clots separately, identifying two of four. The authors considered that SCT is indicated when VP results are indeterminate, and they recommended following up negative or equivocal SCT with pulmonary angiography to detect peripheral emboli.

Goodman *et al.*[87] set out prospectively to compare SCT with selective pulmonary angiography in a patient group with suspected acute PE and indeterminate findings on VP scans, but without deep vein thrombosis or contraindications for intravenous contrast material. Twenty patients completed the study. The average age was 53 years (range 25–84). CT and pulmonary angiograms were compared vessel for vessel, with results grouped per patient and per lung; thus were sensitivity and specificity determined. These values were 63% and 89% when all vessels were included; 86% and 92% for larger, central vessels only. The authors concluded that a major limitation of SCT is poor visualisation of subsegmental vessels, so that a normal CT scan does not exclude the presence of small, subsegmental emboli. They also noted lower interobserver agreement concerning the presence or absence of PE in a given lung for the CT scans than for the pulmonary angiograms. They concluded that, in following up indeterminate VP scans, pulmonary angiography is the procedure of choice.

van Rossum *et al.*[88] also considered SCT as a follow-up test to indeterminate VP. In their study of patients with suspected acute PE, 48 of 67 with indeterminate VP findings were also investigated with both SCT and conventional selective pulmonary angiography. Other patients in the study did not receive the gold standard of pulmonary angiography. In the group of 48, for all vessels, a sensitivity of 80% was recorded for one observer,

TABLE 17 *Studies comparing automated scan technology for liver lesions – study characteristics*

	Kopka *et al.* 1995[83]	**Kopka *et al.* 1996[84]**	**Silverman *et al.* 1995[85]**	**Silverman *et al.* 1996[86]**
Centre	Goettingen, Germany	Goettingen, Germany	Washington, DC	Washington, DC
No. patients	30 30	30 30 30 30	56 53	27 44 29 47 36
Male (%)	45	N/S	43 40	38
Age (years)	41–78 (av. 60) 42–77 (av. 60)	28–78 (av. 60) 32–76 (av. 60) 33–75 (av. 61) 35–78 (av. 61)	Analysed using *t*-test	37–84 (av. 59) 22–83 (av. 59) 37–89 (av. 60) 25–86 (av. 55) 25–87 (av. 57)
Weight (kg)	54–95 (av. 70) 55–91 (av. 70)	58–92 (av. 70) 56–98 (av. 71) 57–95 (av. 71) 55–97 (av. 71)	Analysed using *t*-test	43–100 (av. 66) 45–118 (av. 70) 40–100 (av. 71) 48–105 (av. 71) 50–86 (av. 67)
Scanner	HiSpeed Advantage	HiSpeed Advantage	HiSpeed Advantage	HiSpeed Advantage
Slice (mm)	10	7	N/S	10
Pitch	1	HAP 1.6 PVP 1.3	N/S	N/S
Reconstruction interval (mm)	N/S	3	N/S	N/S
Voltage (kV)	120	N/S	N/S	N/S
Current (mA)	300	220 HAP 280–320 PVP	55	55
Injection rate (ml/s)	4	4	2.5	3
Contrast	Iopromide 300	Iopromide 300	Iohexol 300	Ioversol 320

67% for the other. Both achieved 100% specificity. The authors noted that accuracy in the subsegmental branches was poor, and recommended research into the clinical relevance of subsegmental PE. Their final conclusions related more to the part of their study that compared SCT directly with VP, but without the gold standard of pulmonary angiography that was required for inclusion in this review.

EBCT – CAD

For EBCT, three roles were investigated:

- the prediction of CAD in asymptomatic individuals
- the diagnostic accuracy of angiographic CAD in symptomatic patients
- the use of EBCT to track disease progression, reproducibility.

For the prediction of CAD in asymptomatic patients, only two studies were found that fulfilled the criterion of length of follow-up greater than 1 year (*Table 23*).[89,90] For diagnostic accuracy, 13 studies satisfied the qualitative inclusion criteria;[22,70,91–101] seven of these also satisfied the additional criteria for inclusion in the quantitative analysis (*Table 24*). In addition, four studies were included in this category because, although they did not investigate the diagnosis by EBCT with angiography, they provided valuable information on other aspects of this type of investigation (*Table 25*).[102–105] In the last role, tracking disease progression, nine studies were included:[22,98,106–112]

TABLE 18 *Studies comparing PVP protocols – parameters investigated*

Reference	Group	No. patients	Delay time (s)	Contrast volume (ml)	Injection rate (ml/s)	Contrast medium (mg I/ml)
Brink *et al.* 1995[77]	1	31		125		350
	2	30		100		350
	3	31		75		350
	4	28		125		300
	5	31	60	100	40% at 5	300
	6	32		75	60% at 2	300
	7	29		125		240
	8	28		100		240
	9	7		75		240
	10	30		125		350
	11	28		100		350
	12	30		75		350
	13	30		125		300
	14	30	60	100	3	300
	15	31		75		300
	16	31		125		240
	17	30		100		240
Herts *et al.* 1995[78]	1		45	100		Iohexol 300
	2		45	125		Ioversol 240
	3		45	125		Ioversol 240
	4	24–32	60	100	2	Iohexol 320
	5		60	125		Ioversol 240
	6		60	125		Ioversol 320
Hoeffel *et al.* 1996[79]	1	40	30	120	3	Iohexol 240
	2	15	40	120	2, 3	Iohexol 240
	3	26	50	120	2, 3	Iohexol 240
	4	20	70	105	1.5, 2	Iohexol 300
	5	20	70	125	1.5, 2	Iohexol 300
Kopka *et al.* 1995[80]	1a	25	40		2	
	1b		70			Iopromide
	2	25	40	100	4, 2	300
	3	25	40		4	
Polger *et al.* 1995[68]	1	16	30	75		
	2	15	60	75		
	3	20	40	100	2.5	Diatrizoate
	4	16	60	100		meglumine
	5	25	60	150		
Silverman *et al.* 1995[81]	1	N/S	50	150	3	300
	2	N/S	75	150	3	Non-ionic
Small *et al.* 1994[82]	1	5		75		
	2	5	30	100	3, 4, 5	Iohexol 300
	3	5		125		
	4	5		150		

six addressed interobserver variability, three intra-observer variability, and five the reproducibility of repeat examinations (*Table 26*).

Although most of the studies set out to investigate the accuracy of EBCT when used to diagnose angiographically significant CAD, many turned the question around and also investigated the accuracy of EBCT for demonstrating the absence of such disease.

Agatston *et al.*[22] studied 584 consecutive subjects (aged 30–69 years). This is the article that introduced the scoring technique that is now used

TABLE 19 *Studies comparing automated scan technology for liver lesions – parameters investigated*

Reference	Group	HAP delay time (s)	PVP delay time (s)	Dose (ml)
Kopka *et al.* 1995[83]	1	40	N/P	100
	2	auto	N/P	100
Kopka *et al.* 1996[84]	1	20	45	100
	2	auto	HAP + 10	100
	3	20	50	120
	4	auto	HAP + 15	120
Silverman *et al.* 1995[85]	1	70	N/P	150
	2	auto	N/P	150
Silverman *et al.* 1996[86]	1	60	N/P	100
	2	60	N/P	150
	3	auto	N/P	100
	4	auto	N/P	125
	5	auto	N/P	150

TABLE 20 *Studies comparing PVP protocols – study conclusions*

Reference	Injection rate (ml/s)	Delay time (s)	Uniphasic or biphasic injection	Contrast dose giving optimal enhancement
Brink *et al.* 1995[77]	–	–	Uniphasic better	320 or 300 mg/ml in 125 ml
Herts *et al.* 1995[78]	–	60 better than 45	–	320 mg/ml in 125 ml
Hoeffel *et al.* 1996[79]	Slower rates better	70 better than 50	Biphasic better	300 mg/ml in 125 ml
Kopka *et al.* 1995[80]	4 better than 2	70 better than 40	Uniphasic better	–
Polger *et al.* 1995[68]	–	N/Sig 40 or 70	–	N/Sig
Silverman *et al.* 1995[81]	–	75 better than 50	–	–
Small *et al.* 1994[82]	5 better than 3	–	–	300 mg/ml in 125 ml or 150 ml

–, effect of parameter not investigated; N/Sig, not statistically significant

TABLE 21 *Studies comparing automated scan technology for liver lesions – study conclusions*

Reference	HAP delay time from autoscan (s)	PVP delay time from autoscan (s)	Conclusions
Kopka *et al.* 1995[83]	34–56 (av. 43)	N/P	Automated enhancement 55.1 ± 6.9 HU Standard enhancement 48.4 ± 12.1 HU
Kopka *et al.* 1996[84]	15–33 (av. 21)	38–54 (av. 44)	Automated inadequate HAP timing in 7 compared with 24 for fixed delay Fixed PVP peaked too early in 7 compared with none for automated
Silverman *et al.* 1995[85]	48–86 (av. 69.6)	N/P	94% of autoscan reached 50 HU (mean 71.6 ± 15.2 HU) 66% of fixed scan reached 50 HU (mean 59.8 ± 20.1 HU)
Silverman *et al.* 1996[86]	44–83 (av. 65.8)	N/P	Automated gave better enhancement for 150 ml groups No difference for 100 ml groups

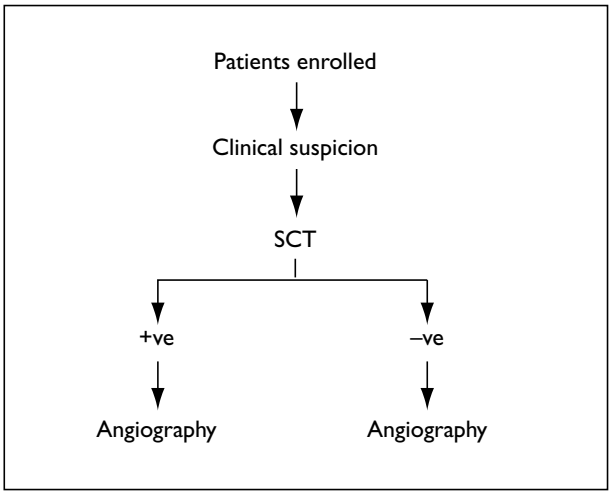


FIGURE 5 *Role of SCT in PE: used in place of VP*

almost universally. There was a mix of patients, including those with established CAD, suspected CAD and asymptomatic individuals. The gold standard diagnosis of clinical CAD was not made purely on angiographic evidence. Of the 109 positive patients, 87 were diagnosed by angiography; the remaining 22 had a documented history of myocardial infarction. Similarly, normal angiographic results were available for 35 of 308 patients in the negative group. There was, however, a subgroup analysis in 50 patients that is relevant to the question posed in this review concerning the comparison of EBCT with angiography. In this group, for a calcium score > 0, the sensitivity and specificity for angiographic obstructive disease were 100% and 28%. The authors based their conclusions on the high NPV found in the total patient group, and suggested that EBCT might be useful for screening.

Barbir *et al.*[91] investigated a group of cardiac transplant recipients whose symptoms included chest pain, so this population is somewhat different from those in the other studies. EBCT calcium score diagnoses were compared with angiographic diagnoses for three angiographic dichotomies. The authors concluded that the test had a potential role as a screening test for postoperative coronary angiography to demonstrate the **absence** of coronary calcification.

Bielak *et al.*[92] studied a population that included 160 routine, symptomatic clinical subjects. Their main purpose was to investigate the repeatability of EBCT, and several sets of results using different dichotomies were included. Sensitivity and specificity results were presented graphically and showed that, although there was a reduction in sensitivity (from 97% to around 80%) if the minimum area classified as disease is increased, the specificity improved markedly from around 50% to about 80%.

Bormann *et al.*[93] studied 50 patients who had also undergone coronary artery angiography. The main thrust of this article was an attempt to correlate calcium with site-specific atherosclerosis, but their comparison of calcium score to stenosis in any proximal vessel is appropriate to this review. A sensitivity of 94% was reported, but with a corresponding specificity of 26%. The authors concluded that the best role might be to use the absence of calcium to rule out significant stenosis.

Braun *et al.*[94] investigated several patient groups. In the 102 who received both EBCT and angiography,

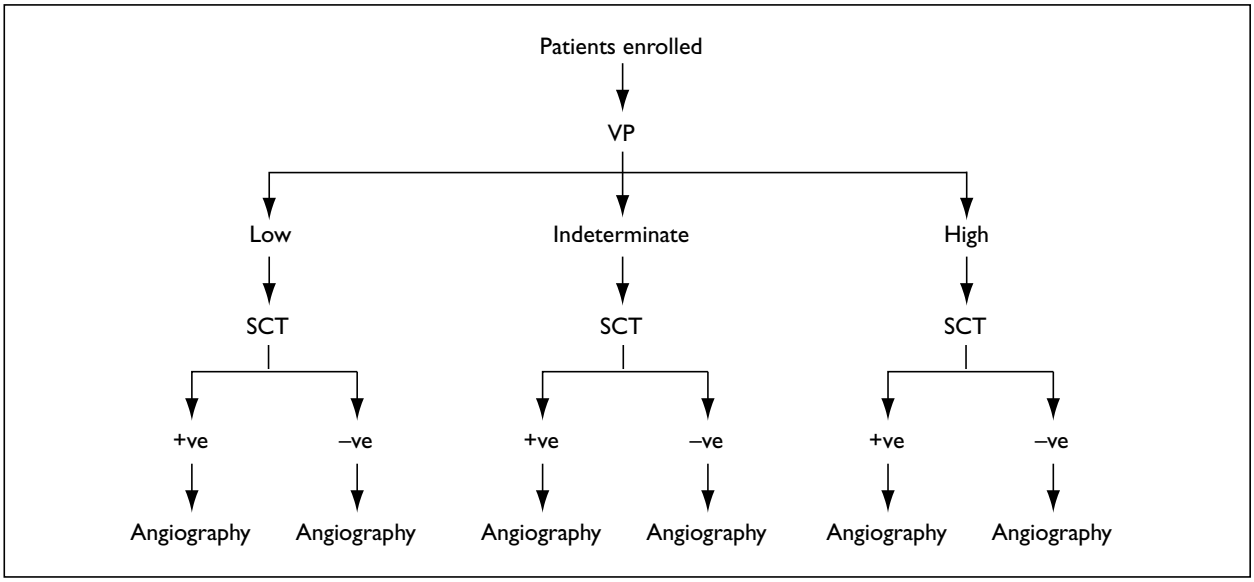


FIGURE 6 *Role of SCT in PE: used together with VP*

TABLE 22 *Details of SCT – PE studies included in the review*

	Goodman et al. 1995[87]	**Remy-Jardin et al. 1996[69]**	**Remy-Jardin et al. 1992[20]**	**van Rossum et al. 1996[88]**
Centre	Milwaukee, WI	Lille, France	Lille, France	The Hague, The Netherlands
No. patients verified with angiography	20	75	42	56
Age range (mean) (years)	25–84 (53)	22–83 (59)	21–65 (34)	N/S
Male (%)	60	43	71	N/S
Patient group	Indeterminate or inconsistent VP	Clinical suspicion	Clinical suspicion	Indeterminate or inconsistent VP
No. emboli	N/S	195	112	N/S
Location of emboli	Main Lobar Segmental Subsegmental	Main Lobar Segmental	8 main 28 lobar 76 segmental	Includes subsegmental
SCT coverage	12 cm from aortic knob	Central 10 cm excluding branches with poor vascular opacification	Central 12 cm excluding branches with poor vascular opacification	16 cm from aortic knob
Angiographic coverage	Left and right main + additional oblique	Unilateral, lobar, segmental, right and left posterior oblique or frontal	Unilateral, right and left posterior oblique	Anteroposterior and 30° degree right and left anterior oblique
SCT standard(s)	Partial or complete filling	Vascular PE signs as reported in literature	Partial filling defects Complete filling defects Railway track signs Mural defects	Partial or complete intraluminal fillings
Angiographic standard(s)	Vessel cut-off with meniscus or contrast material tracking around an intraluminal thrombus	Intraluminal filling defect Secondary signs of PE	Intraluminal filling defect Secondary signs of PE	Intraluminal filling defect Secondary signs of PE
Conclusions	4 of 11 had emboli in subsegmental vessels only CT showed only 1 of these 4	Agreement of 188 emboli 10 emboli on SCT only 7 emboli on angiography only	Agreement of 112 emboli in 18 patients 10 disagreements (false-positives on CT) in 4 of these 18 patients	SCT poor for clots beyond segmental arteries

EBCT had a sensitivity of 93% and selectivity/specificity of 73%. The authors were interested in calcium rather than angiographic disease and so drew no conclusions about the test in that role.

Breen *et al.*[95] studied 100 patients who were under 60 years of age, who had clinical indications for, and underwent, coronary angiography. For the diagnosis of clinically significant disease, sensitivity and specificity were 100% and 47%, and PPV and NPV values were 63% and 100%. The authors emphasised the high NPV in this young population, and suggested that false-positives may have been so classified because of the inadequacies of angiography in identifying atherosclerosis.

Budoff *et al.*[96] presented a multicentre study, with some of the results being reported separately in other articles. The resulting study population included 710 patients, with several characteristics of the population differing between sites. The overall sensitivity and specificity were 95% and

TABLE 23 *Studies included for prediction of CAD in asymptomatic individuals*

Reference	Centre	No. patients	Male (%)	Age (years)	EBCT protocol	Patient dose (mGy)	EBCT dichotomy (CAC score unless otherwise stated)	Follow-up	Subgroup analyses	Conclusions
Arad *et al.* 1996[89]	Roslyn, NY	1173	71	53 ± 11	40 slices at 3 mm	N/S	100 160 680	19 months (99.8%)	None	Only hypertension correlated with CAC score Only age correlated with subsequent cardiovascular events
Coin 1993[90]	Miami, FL	147	N/S	N/S	N/S	N/S	Coronary score not calculated	5 years	None	Of 147 asymptomatic patients, 43 showed coronary disease, confirmed in 41 by unspecified clinical and laboratory findings Stated that clinical correlation maintained at 5 years, but no data

TABLE 24 *Accuracy studies for diagnosis of CAD in symptomatic individuals, using the EBCT calcium score*

Reference	Centre	No. patients	Male (%)	Age (years)	EBCT protocol	Patient dose (mGy)	EBCT dichotomy (CAC score unless otherwise stated)	Angiography dichotomy	Sub-group analyses	Conclusions
Agatston *et al.* 1990[22]	Miami Beach, FL	584	70	48 ± 10	20 slices at 3 mm	< 5	1, 25, 50, 100, 200, 300, 500, 700	50% diameter	No. slices Age	A score of 50 appeared to be the best for age 40–50, and a score of 300 for those aged 60 Extra slices found calcium in one patient exclusively
Barbir *et al.* 1994[91]	London, UK	102	86	Men: 53.5 Women: 51.2	40 slices at 3 mm	N/S	0	25% 50% 75% diameter	Vessel	45% of cardiac transplant patients had coronary calcification after a median of 4.6 years Potential role to exclude CAD
Bielak *et al.* 1994[92]	Rochester, MN	160	83	48.5 ± 7.6	40 slices at 3 mm	10	2, 4, 6, 8, 10, 12 pixels	10% 50% diameter	None	Sensitivity increases as the number of pixels is reduced, but corresponding reduction in specificity Sensitivity is higher for severe disease
Bormann *et al.* 1992[93]	Iowa City, IA	50	53	57.9 (range 37–82)	20 slices at 3 mm	N/S	0	50% diameter	Vessel Age	Older (> 60 years) patients were found to have higher mean calcium scores (p = 0.005)
Braun *et al.* 1996[94]	Nürnberg, Germany	102	72	56 ± 8	30 slices at 3 mm	N/S	N/S	50% (diameter/ area not stated)	Age	Median calcium score and range per decade of age Age is significantly related to calcium score for both normal and positive patients

continued

TABLE 24 contd *Accuracy studies for diagnosis of CAD in symptomatic individuals, using the EBCT calcium score*

Reference	Centre	No. patients	Male (%)	Age (years)	EBCT protocol	Patient dose (mGy)	EBCT dichotomy (CAC score unless otherwise stated)	Angiography dichotomy	Sub-group analyses	Conclusions
Breen *et al.* 1992[95]	Rochester, MN	100	91	47.1 ± 7.6	40 slices at 3 mm	8.5	0	0% diameter 1–49% diameter 50% diameter	None	Patients with concordant findings from EBCT and angiography had a higher mean age than those with discordant findings
Budoff *et al.* 1996[96]	Multicentre	710	64	56 ± 12	20–40 slices at 3 mm	10	0	50% diameter	Age Vessel Sex ln score	Increasing sensitivity, decreasing specificity for CAD prediction with age No difference for sensitivity, but higher specificity in women ($p < 0.001$)
Devries *et al.* 1995[97]	Chigaco, IL	140	50	Men: 56 ± 12 Women: 60 ± 12	40 slices at 3 mm	N/S	0	0% diameter 70% diameter	Age Sex Vessel	Sensitivity for any CAD (> 0%) for women < 60 years is less than that for men < 60 and for women > 60 ($p < 0.005$) Sensitivity for 70% CAD is no different
Fallavollita *et al.* 1994[98]	Buffalo, NY	106	74	43.6 ± 5.4 (91% Caucasian)	20 slices at 3 mm	5	1, 2, 4 voxels	0% diameter 1–49% diameter 50% diameter	Vessel Observer	Calcium presence was related to older ($p = 0.002$), male ($p = 0.001$) and hypercholesterolaemic ($p = 0.015$) patients
Kajinami *et al.* 1995[70]	Kanazawa, Japan	251	69	56 ± 14	20 slices at 3 mm	N/S	ln score = 3.0 ln score per decade	75% densitometric narrowing or 50% diameter	Age	Higher ln(calcium score) for patients with CAC both for total group and by age Difference most prominent in men 40–60 years and women > 60 years
Rumberger *et al.* 1995[99]	Rochester, MN	139	64	Men: 47 ± 7 Women: 56 ± 11 (100% Caucasian)	40 slices at 3 mm	N/S	0	0% diameter 1–20% diameter 21–49% diameter 50% diameter	Sex	No difference between men and women for the detection of calcium by EBCT
Tanenbaum *et al.* 1989[100]	Chigaco, IL	54	67	54 ± 16	8 mm 4 mm gap	N/S	0	50% diameter in left main artery, 70% diameter elsewhere	None	Sensitivity increased for more extensive disease
Yaghoubi *et al.* 1995[101]	Torrance, CA	67	48	55 (range 30–75)	20–40 slices at 3 mm	10	0	50% diameter	Observer Scoring system	Alternative calcium analysis system has equivalent accuracy to Imatron Subsequent time and storage advantages

TABLE 25 *EBCT studies of CAC that were excluded from the diagnostic performance review but which address relevant issues and are discussed in chapter 8*

Reference	Centre	No. patients	Male (%)	Age (years)	EBCT protocol	Patient dose (mGy)	EBCT dichotomy (CAC score unless otherwise stated)	Angiography dichotomy analyses	Sub-group	Conclusions
Agatston *et al.* 1994[102]	Miami Beach, FL	100	76	58 ± 11	20 slices at 3 mm	5	Not calculated	0 > 5% area > 75% area	None	Square root of plaque volume and square root of number of lesions related to number of segments
Kaufmann *et al.* 1995[103]	Rochester, MN	145	83	20–59 (100% Caucasian)	40 slices at 3 mm	10	0	10% diameter increments	Age Sex	Prevalence of calcium and score range
Shields *et al.* 1996[104]	Spokane, WA	104	65	53 (range 27–79)	40 slices at 3 mm	N/S	1	N/S	No. slices	59 of 104 positive for total scan; 53 positive for 10 slices; 57 for 20 slices
Wong *et al.* 1994[105]	Torrance, CA	1218	76	53 ± 10	20–30 slices at 3 mm	N/S	0, 50, 100, 250, 500	N/S	Age Sex ln score	Only age and ln(score) were significant for CAC score as an indicator of CAD ($p < 0.01$)

44%, the PPV and NPV being 72% and 84% respectively. The authors considered the most promising applications to be: (1) in patients over 50 years of age, in whom a CAC score of zero excludes significant stenosis; and (2) in a population less than 40 years of age in whom a CAC score greater than zero diagnoses stenosis. A reanalysis was undertaken to attempt to improve the specificity of the test by using the number of vessels affected rather than the CAC score. An improvement in specificity was at the expense of sensitivity, but the inclusion of other risk factors in the model improved results. A comprehensive discussion of the limitations of the study was included.

Devries *et al.*[97] studied 140 patients who underwent coronary arteriography. The diagnosis of CAD included both lumen irregularity and stenosis ≥ 70%. For obstructive disease, the overall sensitivity and specificity were 97% and 41%, and PPV and NPV 55% and 94%. For any lumen irregularity, the overall sensitivity and specificity were 88% and 55%; PPV and NPV were 82% and 66%. The authors concluded that the test has a strong NPV for detecting obstructive coronary disease.

Fallovollita *et al.*[98] were interested in a young population, in whom the specificity of the test may be higher than for older groups. They looked at 106 patients aged under 50 years. The overall sensitivity and specificity were 85% and 45%; the PPV and NPV were 66% and 70%. They concluded that both the specificity and the NPV are inadequate in the under 50 years age group.

Kajinami *et al.*[70] designed a study to compare EBCT with ECG and thallium exercise tests. All 251 patients underwent coronary angiography. They investigated age and sex variations and ROC curve analysis to determine threshold CAC scores for the diagnosis of stenosis. Having optimised this threshold, their overall results for sensitivity and specificity were 77% and 86%, with PPV and NPV being 86% and 76%. They concluded that EBCT is a useful test for predicting angiographically-defined coronary atherosclerosis. This was the only study to use densitometrically-defined angiographic stenosis.

Rumberger *et al.*[99] investigated 139 patients. Their results were presented separately for men and women. For men, the sensitivity and specificity were 94% and 35%, and the PPV and NPV were 89% and 79%. For women, the sensitivity and

TABLE 26 *EBCT CAC studies addressing observer or interexamination reproducibility*

Reference	Centre	No. patients	Age (years)	EBCT protocol	Calcium or CAD	Assessment	Time between scans/reading	Statistical method	Conclusions
Agatston *et al.* 1990[22]	Miami Beach, FL	88	N/S	20 slices at 3 mm	Calcium	Interobserver	N/S	None	70 of 88 scored identically
Devries *et al.* 1995[106]	Chigaco, IL	91	55 ± 12	40 slices at 3 mm	Both	Interexamination	Within 24 h	Difference/ mean	72 ± 54% for < 10 score; 60 ± 46% for 11–100 score; 28 ± 26% for > 100 score 11% for > 0 CAC prediction of CAD; 9% for > 2 CAC; 5% for > 4 CAC
Fallavollita *et al.* 1994[98]	Buffalo, NY	106	25–49	20 slices at 3 mm	Both	Intraobserver Interobserver	N/S	None	103 of 106 agreed CAC presence/absence Agreement in 391 of 424 vessels Scores agreed by ≤ 1 in 68, av. difference 16%
Hernigou *et al.* 1996[107]	Paris, France	41 59 50	51 ± 5.8 47.8 ± 8.8 51.5 ± 8.5	20 slices at 3 mm 4 pixels	Calcium	Intraobserver Interobserver Interexamination	Within 24 h	(mean variance)$^{1}/_{2}$/ mean	4.7% intraobserver reproducibility 4.5% interobserver reproducibility 22% interexamination reproducibility
Kajinami *et al.* 1993[108]	Kanazawa, Japan	75	57 ± 14	20 slices at 3 mm	Calcium	Interobserver Intraobserver Interexamination	Several min	Bland and Altman[113]	No difference between intraobserver 57 out of 75 scored identically Mean difference from scans = 1.8 ± 106
Kaufmann *et al.* 1994[109]	Rochester, MN	25	23–59	30–40 slices at 3 mm 2 pixels	Calcium	Intraobserver Interobserver	5 months	ANOVA reliability coefficient	Intraobserver good regardless of CAC score High coefficients for 3 observers for score, lesions and area
Shields *et al.* 1995[110]	Spokane, WA	50	Men: 52.3 Women: 56.6	20 slices at 3 mm	Both	Interexamination	12 min	Matched pairs *t*-test Cronbach's alpha	Only total number of lesions differed Reliability coefficient 0.99 for lesion number, volume and score 8 discordant CAD diagnoses
Shields *et al.* 1996[111]	Spokane, WA	50	53.9	20 slices at 3 mm	Both	Interobserver	12 min	Inter-rater *k* Pearson's product-moment coefficient	Only 1 discrepancy from observers using CAC > 10 for CAD Agreement was reduced by using > 0 High reliability for score, volume and number
Wang *et al.* 1996[112]	Torrance, CA	175 72 77	63 ± 8	20 slices at 3 mm 30 slices at 3 mm 20 slices at 6 mm	Calcium	Interexamination	2–3 min	Means of pairs into terciles Bland and Altman[113]	6 mm protocol was more reproducible for CAC score, volume and mass ($p < 0.005$) 20 slices at 3 mm = 34% for CAC score

ANOVA, analysis of variance

specificity were 97% and 38%, and the PPV and NPV were 85% and 91%. The authors concluded that there was no significant difference between the performance of the test in men and women, and that the test has a strong NPV for detecting obstructive coronary disease.

In 1989, Tanenbaum *et al.*[100] investigated 54 patients in the very first study of its kind. The overall sensitivity was 88%. Sensitivity and specificity were defined incorrectly in the article; however sufficient information was provided to calculate the correct values of sensitivity 88% and specificity 55%. The authors drew no firm conclusions from their results, other than to recommend further studies.

Yaghoubi *et al.*[101] presented the results from a group of 67 symptomatic patients who were identified in the course of a comparison of analysis systems. For the standard system used also by other investigators, sensitivity was 97% and specificity 56%. The authors drew no conclusions about the role of EBCT.

The four articles in *Table 25* have been included because, although they do not compare the diagnoses made with EBCT and coronary angiography, they investigate other features of this modality that are important for understanding the results of the main set of included studies.[102–105] They are discussed in chapter 8.

The nine studies[22,98,106–112] addressing the role for EBCT of tracking disease progression are discussed next.

Agatston *et al.*[22] investigated interobserver reproducibility. An Imatron C100 was used to acquire ECG triggered, breath-hold studies on a group of 584 patients with and without clinical CAD. Two physicians independently scored 88 of these. Reproducibility was discussed in the article in terms of agreement between observers: 80% of 88 studies were scored identically. Sources of disagreement were described, but no conclusions drawn.

Devries *et al.*[106] investigated interexamination reproducibility. An Imatron C100 was used to acquire two ECG triggered, breath-hold studies within 24 hours on a group of 91 consecutive patients with and without chest pain. One experienced reviewer read the scans in random order and blinded to previous results. Reproducibility was expressed in terms of variability, being the percentage value of difference/mean. A number of parameters were evaluated, including positive/negative diagnosis. They found that a patient with an examination interpreted as positive or negative had a significant likelihood of a discordant finding on a repeat scan. The variability was found to be related inversely to the score, being greater for the lower values. For those with a non-zero calcium score on both scans, variability was 49 ± 45%. It was concluded that changes in scores of several hundred would be required to be sure that true changes in disease state had occurred.

Fallavollita *et al.*[98] investigated intra- and interobserver reproducibility. An Imatron C100 was used to acquire ECG triggered, breath-hold studies on a group of 108 patients, who also underwent coronary angiography for clinical evaluation of CAD. One hundred and six technically adequate studies were obtained. Each study was interpreted by two independent observers. Reproducibility was discussed in the article in terms of agreement between observers, in both the presence and the absence of calcification and in terms of the calcium score: 68 of 106 (64%) had scores that differed by ≤ 1. A regression analysis gave $r^2 = 0.87$ and the average disagreement in scores between observers as 16%. Intraobserver variability was determined on 29 scans, giving a mean difference in calcium scores of 8%. Sources of disagreement were described, but no conclusion was drawn.

Hernigou *et al.*[107] investigated interexamination and intra- and interobserver reproducibility. An Imatron C100 was used to carry out ECG triggered, breath-hold studies on a group of 150 asymptomatic subjects. Different groups were used for the three comparisons; 50 subjects had two studies carried out on the same day. Variability was defined as the square root of the mean of the intrameasurement variance, and a parameter R was given by variability/mean. Only calcium scores ≥ 5 were included in the analysis. Log_{10} conversion was applied to the scores and measurements were made on 12 slices from the slice including the left coronary artery ostium. The interexamination R was 22% for the linear score and 7% for the logarithmic score on 12 slices. Corresponding intraobserver R-values were 4.7% and 1.9%; interobserver values were 4.5% and 1.3%. It was concluded that these values were good enough for EBCT to be used for longitudinal studies.

Kajinami *et al.*[108] investigated interexamination and intra- and interobserver reproducibility. An

Imatron C100 was used to acquire two ECG triggered, breath-hold studies within several minutes on a group of 75 consecutive patients. Two blinded observers were used. Measurements for the intraobserver evaluation were performed several days and 1 year apart. Both total calcium score and ln(1 + total calcium score) were assessed. The Bland and Altman[113] method was used to compare the pairs of measurements. Intraobserver variability was excellent, interobserver reproducibility good and interexamination reproducibility inadequate. Averaging of the results from two examinations was recommended and also the use of the log transformed score, because the differences did not increase with the mean.

Kaufmann *et al.*[109] investigated intra- and interobserver reproducibility. An Imatron C100 was used to carry out ECG triggered studies on a group of 25 patients with CAC. Three blinded observers were used. Measurements for the intraobserver evaluation were performed 5 months apart. Analysis of variance was used. Total scores, number of lesions and area of calcification were assessed. Intraobserver agreement was excellent; interobserver agreement was graded as good or excellent.

Shields *et al.*[110] investigated interexamination reproducibility. An Imatron C100 was used to carry out two ECG triggered breath-hold studies, 12 minutes apart, on a group of 50 consecutive patients. Different radiographers performed the two scans and performed blinded reading of both of them. A paired *t*-test between the mean of the two results and the second result was performed, and the test-retest reliability assessed in terms of Cronbach's alpha procedure. The authors found high reproducibility, but did note poor agreement when the calcium score was low. They recommended the use of EBCT for longitudinal studies.

Shields *et al.*[111] investigated interobserver reproducibility. An Imatron C100 was used to carry out two ECG triggered breath-hold studies, 12 minutes apart, on a group of 50 consecutive patients. Different radiographers performed the two scans and performed blinded reading of both of them. These are the same data as used in the study described above.[110] Several measures of reliability were used, including the kappa statistic. High interobserver agreement was noted.

Wang *et al.*[112] investigated interexamination and interprotocol reproducibility. An Imatron C100 was used to carry out ECG triggered breath-hold studies on a group of 324 consecutive asymptomatic volunteers. The examinations were carried out 2–3 minutes apart. The Bland and Altman[113] methodology and nonparametric Wilcoxon tests were used to assess reproducibility. The authors calculated that their standard protocol (20×3 mm slices) would require a 40–60% change in calcium score to be sure that the change was real. The best interexamination reproducibility was achieved with 20×6 mm slices that still had a difference/mean of 14% for calcium score. The conclusion was that EBCT measurement of calcification is not sufficiently reliable to assess serial changes in calcification in individual patients, but that group differences over time may be assessed.

Radiation dose

There are three commonly used measures of radiation dose in CT. Peak radiation dose gives the peak dose from a single scan. More clinically relevant is the multiple scan average dose (MSAD), which gives the maximum dose to the image volume when multiple scans are acquired. The CT dose index (CTDI) is defined as the integral of a single scan dose profile along an infinite line perpendicular to the slice plane, divided by the nominal slice thickness. CTDI can be used to estimate the MSAD; it is measured over 10 cm or over 14 slice thicknesses. It can result in errors if the radiation dose profile has a full width, half maximum value that differs from the nominal scan thickness. In addition, the exposure-to-dose conversion used is often for acrylic rather than muscle or water, which leads to underestimation.

SCT

Tables 27–30 detail the dosimetry techniques, the technical details of comparative modalities, the study results and the study conclusions respectively for the eight studies included.[114–121] Four studies compared the radiation dose of SCT with that of conventional CT, ensuring that parameters such as slice thickness and tube voltage and current were similar between modalities.

EBCT

The literature on radiation dosimetry is less comprehensive for EBCT. Three articles[122–124] were identified. Two of these described a methodology for measurement that allows for the complex dose distribution of EBCT. The method was applied in the third article to determine the dose to the

breast from chest, coronary artery and cardiac function EBCT examinations (*Table 31*).[124] Values from conventional CT were also presented. The authors concluded that, in spite of differences in dose distribution, the dose from EBCT was comparable with that from conventional CT.

Radiation doses were also quoted in eight of the articles included in the 'Diagnostic performance' section of this review. It was not possible to evaluate the experimental methodology for these measurements, so all are included in *Table 32* for information.[22,92,95,96,98,101–103]

TABLE 27 *SCT dose – dosimetry techniques*

Reference	Examination	Dosimeter	Subject	Placement
Chang *et al.* 1995[114]	Brain Chest Abdomen	LiF TLD-100 chips	Rando phantom	Sandwiched into a close contact slice
Collie *et al.* 1994[115]	Pulmonary metastases	LiF TLD pellets	Rando phantom	6 pellets at 3 sites in slice
Jurik *et al.* 1996[116]	Sternoclavicular joints	LiF TLD chips	Rando phantom	4 packets in 27 holes 4 x 3 points – skin
Jurik *et al.* 1996[117]	Pelvis	LiF TLD packages	Rando phantom	4 packets in 24 holes 4 x 2 points – skin
Liang *et al.* 1996[118]	N/App	Pencil ionisation chamber	FDA CTDI body phantom	5 holes and surface Lucite rods used
Lucidarme *et al.* 1996[119]	Bronchi	LiB TLD chips	Plexiglass cylinder: 2 holes	Series of 10 TLDs in centre of each hole
McGhee and Humphries 1994[120]	N/App	LiF TLD-100 ribbons	Acrylic body and head phantoms	150 TLDs in 1 central hole + 4 peripheral
Verdun *et al.* 1996[121]	N/App	Pencil ionisation chamber	No phantom present during dose measurement	Centre of rotation of scanner

LiF, lithium flouride; TLD, thermoluminescent dosimeter; N/App, not applicable; FDA, Federal Drug Administration; CTDI, computed tomography dose index; LiB, lithium borate

TABLE 28 *SCT dose – technical details for the comparative modalities*

Reference	Modality	Scanner	Slice	Pitch	Voltage, exposure	Slices/ coverage	Additional scan(s)
Chang *et al.* 1995[114]	SCT	GE Prospeed	Brain: 7–8 mm Chest: 10 mm Abdomen: 10 mm	1	120 kV, 200 mA	15 x 2 23 x 2 21 x 2	120 kV, 80 mA
	CCT	GE Prospeed	Brain: 7–8 mm Chest: 10 mm Abdomen: 10 mm	N/App	120 kV, 160 mA	15 x 2 23 x 2 21 x 2	120 kV, 80 mA
Collie *et al.* 1994[115]	SCT	Somatom Plus	10 mm	1	137 kV, 145 mA	24	None
	CCT	Somatom Plus	10 mm	N/App	137 kV, 145 mA	24	None
Jurik *et al.* 1996[116]	SCT	Somatom Plus S	3 mm	N/S	120 kV, 210 mAs	9 cm	120 kV, 340 mAs
	CCT	Philips BT-S4	8° elliptical at 1 cm intervals	N/App	68 kV, 60 mA	7 for 10 x 9 cm	18 x 8 cm 65 kV 100 mA
Jurik *et al.* 1996[117]	SCT	Somatom Plus S	3 mm	N/S	120 kV, 210 mAs	18 cm	120 kV, 340 mAs
	Radiography	N/App	5 radiographs and 19 x 23 cm	N/App	65 kV, 184 mA +2 views and 70 kV, 150 mA, 70 kV, 212 mA	N/App	None
Liang *et al.* 1996[118]	Single SCT	GE HiSpeed	5 mm	1	120 kV	45 mm	None
	Dual SCT	Elscinct Twin	2 x 5 mm	2	120 kV	45 mm	None
Lucidarme *et al.* 1996[119]	SCT	Tomoscan SR7000	3 mm	1.6	120 kV, 150 mAs	57 and 2 mm overlap	None
	Thin section	Tomoscan SR7000	1.5 mm with 10 mm intervals	N/App	120 kV, 175 mAs	N/S	None
McGhee and Humphries 1994[120]	SCT	Somatom Plus	5 or 10 mm	1	120 kV, 500 mA	120 mm	None
	CCT	Somatom Plus	10 mm; 10 mm interslice gap	N/App	120 kV, 500 mA	11	None
Verdun *et al.* 1996[121]	SCT	GE HiSpeed	7 mm	1	120 kV, 320 mA	5	None
	CCT	GE HiSpeed	7 mm	N/App	120 kV, 320 mA	Same volume	None

SCT, spiral computed tomography; CCT, conventional computed tomography

TABLE 29 *SCT dose – results of included studies*

Reference	Modality	Normalised exposure dose (mR/mAs)	Absorbed dose (mGy)	Effective dose (mSv)	Effective dose equivalent (mSv)
Chang *et al.* 1995[114]	SCT	N/S	N/S	Brain: 1.89 Chest: 30.01 Abdomen: 12.85	Brain: 0.42 Chest: 16.72 Abdomen: 10.33
	CCT	N/S	N/S	Brain: 4.95 Chest: 40.65 Abdomen: 19.62	Brain: 1.65 Chest: 19.98 Abdomen: 12.75
Collie *et al.* 1994[115]	SCT	N/S	Lung: 12.7 Thyroid: 5.9 Breasts: 11.2 Ovaries: 0.1	N/S	Total: 6.5
	CCT	N/S	Lung: 12.8 Thyroid: 7.6 Breasts: 13.3 Ovaries: 0.1	N/S	Total: 7.0
Jurik *et al.* 1996[116]	SCT	N/S	Thyroid: 3.75 Marrow: 6.66 Lung: 7.40 Oesophagus: 5.78 Skin: 26.75	Thyroid: 0.05 Marrow: 0.12 Lung: 0.36 Oesophagus: 0.07 Skin: 0.02 Total: 0.62	N/S
	CCT	N/S	Thyroid: 10.13 Marrow: 11.75 Lung: 6.83 Oesophagus: 19.1 Skin: 164.83	Thyroid: 0.25 Marrow: 0.14 Lung: 0.16 Oesophagus: 0.24 Skin: 0.03 Total: 0.82	N/S
Jurik *et al.* 1996[117]	SCT	N/S	Colon: 10.95 Marrow: 10.78 Bladder: 12.95 Ovaries: 11.79 Skin: 19.14	Colon: 0.92 Marrow: 0.45 Bladder: 0.65 Ovaries: 2.36 Skin: 0.03 Total: 4.41	N/S
	Radiography	N/S	Colon: 12.12 Marrow: 12.21 Bladder: 19.26 Ovaries: 11.36 Skin: 56.45	Colon: 1.16 Marrow: 0.59 Bladder: 0.96 Ovaries: 2.27 Skin: 0.06 Total: 5.04	N/S
Liang *et al.* 1996[118]	Single SCT	Skin: 27.20	N/S	N/S	N/S
	Dual SCT	Skin: 28.84	N/S	N/S	N/S
Lucidarme *et al.* 1996[119]	SCT	N/S	Central: 7.0–8.0 Peripheral: 8.0–15.0	N/S	N/S
	Thin section	N/S	Central: 2.0–4.5 Peripheral: 2.0–13.0	N/S	N/S
McGhee and Humphries 1994[120]	SCT Head	N/S	Centre: 4.21 Periphery: 4.59	N/S	N/S
	SCT Body	N/S	Centre: 1.69 Periphery: 3.89	N/S	N/S
	CCT Head	N/S	Centre: 4.23 Periphery: 4.61	N/S	N/S
	CCT Body	N/S	Centre: 1.68 Periphery: 3.97	N/S	N/S
Verdun *et al.* 1996[121]	SCT	N/S	62	N/S	N/S
	CCT	N/S	62	N/S	N/S

TABLE 30 *SCT dose – conclusions of the included studies*

Reference	Conclusions
Chang *et al.* 1995[114]	CCT has a higher effective dose than SCT
Collie *et al.* 1994[115]	Radiation dose is similar for spiral and CCT Increasing pitch did not significantly reduce resolution
Jurik *et al.* 1996[116]	CCT had a greater dose than SCT for all organs except lung The total effective dose was 0.2 mSv lower with SCT
Jurik *et al.* 1996[117]	Conventional 5-view radiography had a greater dose than SCT for all organs except the ovaries The total effective dose was 0.6 mSv lower with SCT
Liang *et al.* 1996[118]	Dual slice mode has a similar exposure dose to that of single slice mode for the same slice width, averaged per slice and normalised for tube current (mR/mAs) Exposure distribution along the phantom surface was almost isotropic
Lucidarme *et al.* 1996[119]	The total dose delivered to the skin with helical CT was 3.4 times greater than that delivered with thin-section CT The peak skin exposure was the same for both techniques
McGhee and Humphries 1994[120]	No significant increase in dose results from the use of SCT, apart from that arising from the additional half rotation required at each end of the volume MSAD can be applied directly to dose estimation for SCT
Verdun *et al.* 1996[121]	SCT and CCT are equivalent for dose

TABLE 31 *Clinical dose to breast for various EBCT protocols compared with CCT*

Reference	Centre	Dosimetry method	EBCT protocol			Exposure–dose conversion	EBCT absorbed dose			Conventional chest absorbed dose (mGy)
			Chest	CAC	Cardiac function		Chest (mGy)	CAC (mGy)	Cardiac function (mGy)	
McCollough *et al.* 1995[123]	Rochester, MN	Film	40 slices at 6 mm 130 kV, 250 mA 100 ms exposure	40 slices at 3 mm 130 kV, 62.5 mA 100 ms exposure	8 slices at 8 mm 130 kV, 31.3 mA 100 ms exposure	8.15 mGy/R conversion factor 38.8	21.9	2.9	1.6	17.8–32.9 (various scanners)

TABLE 32 *Radiation dose during EBCT examination for CAD in symptomatic patients*

Reference	Centre	Patient dose (mGy)	Organ
Agatston *et al.* 1990[22]	Miami Beach, FL	< 5	N/S
Agatston *et al.* 1994[102]	Miami Beach, FL	5	N/S
Bielak *et al.* 1994[92]	Rochester, MN	10	N/S
Breen *et al.* 1992[95]	Rochester, MN	8.5	Skin
Budoff *et al.* 1996[96]	Multicentre	10	Per patient
Fallavollita *et al.* 1994[98]	Buffalo, NY	5	Per patient
Kaufmann *et al.* 1995[103]	Rochester, MN	10	N/S
Yaghoubi *et al.* 1995[101]	Torrance, CA	10	Skin entry dose

Chapter 5

Details of studies excluded from the review

In this chapter the articles excluded from the review are listed, together with reasons for their exclusion. Because the inclusion criteria differ for each category of study, some articles that have been included earlier in one category may also appear here as excluded from a different category. This is particularly true at the health economics level, where six studies[17,58,77,79,82,85] included at other levels are listed here as exclusions. Similarly, articles included for the qualitative review of the diagnostic performance of EBCT have been excluded from the quantitative review because they do not satisfy the more stringent inclusion criteria.

Health economics

Studies were excluded if they did not contain any economic data, or if there were only brief comments or suggestions for cost-effectiveness. The excluded studies[17,58,77,79,82,85,125–149] are not listed separately here because the same reason for exclusion applied to all: none had any relevance to the topic area.

Patient outcome and therapeutic impact

SCT

Exclusions, and reason for exclusion are listed in *Table 33*.[130,131,150–163]

EBCT

Exclusions, and reason for exclusion are listed in *Table 34*.[148,164–167]

Diagnostic impact

SCT – liver lesions

Two studies were excluded because they duplicated results in a later article.[168,169]

SCT – PE

Exclusions and the reasons for exclusion are listed in *Table 35*.[87,88,170–173]

EBCT – CAD

One non-human study was excluded.[174]

Diagnostic performance

Tables 36–39 list the reasons for excluding diagnostic performance studies that fulfilled all preliminary inclusion criteria and were relevant to the specific clinical application.

SCT – liver lesions

Exclusions and the reasons for exclusion are listed in *Table 36*.[64,168,175–182] No studies satisfied the quantitative inclusion criteria because all had inadequate gold standards.

SCT – PE

Exclusions and the reasons for exclusion are listed in *Table 37*.[173,183,184]

EBCT – CAD

Exclusions and the reasons for exclusion are listed in *Table 38*.[174,185–207]

Exclusions and the reasons for exclusion from the quantititive meta-analysis are listed in *Table 39*.[22,70,92,94,97,100,102–105]

Radiation dose

SCT

Exclusions and the reasons for exclusion are listed in *Table 40*.[2,128,180,208–229]

EBCT

Exclusions and the reasons for exclusion are listed in *Table 41*.[230,231]

TABLE 33 *The 16 excluded SCT patient outcome and therapeutic impact studies*

Reference	≤ 10 patients	Non-human	No link made to outcome or therapeutic impact	Not a comparative study
Adachi *et al.* 1994[150]				✔
Engeler *et al.* 1992[130]			✔	
Gavant *et al.* 1995[131]				✔
Gerber *et al.* 1991[151]	✔			
Kauczor *et al.* 1994[152]			✔	
LoCicero *et al.* 1996[153]			✔	
McEnery *et al.* 1994[154]		✔		
Nunez *et al.* 1996[155]	✔			
Padhani *et al.* 1995[156]				✔
Quillin *et al.* 1996[157]				✔
Quint *et al.* 1996[158]			✔	
Roos *et al.* 1995[159]			✔	
Sagy *et al.* 1996[160]	✔			
Tecce *et al.* 1995[161]	✔			
White *et al.* 1996[162]	✔			
Zimmerman *et al.* 1992[163]				✔
Total	5	1	5	5

TABLE 34 *The five excluded EBCT patient outcome and therapeutic impact studies*

Reference	≤ 10 patients	No link made to outcome or therapeutic impact
Anderson *et al.* 1996[164]		✔
Chareonthaitawee *et al.* 1995[165]		✔
Kao *et al.* 1995[166]	✔	
Kimura *et al.* 1990[167]	✔	
Stern *et al.* 1994[148]	✔	
Total	3	2

TABLE 35 *The six excluded SCT PE diagnostic impact studies*

Reference	≤ 10 patients	No reference standard	Non-human study
Dresel *et al.* 1995[170]		✔	
Goodman *et al.* 1995[87]		✔	
Sostman *et al.* 1996[171]	✔		
Steiner *et al.* 1996[172]		✔	
van Rossum *et al.* 1996[88]		✔	
Woodard *et al.* 1995[173]			✔
Total	1	4	1

TABLE 36 *The ten excluded SCT liver lesion diagnostic performance studies*

Reference	≤ 10 patients	Duplicate	No comparative protocol used
Baron *et al.* 1996[175]		✔	
Frederick *et al.* 1996[176]			✔
Irie *et al.* 1994[177]			✔
Irie *et al.* 1996[178]			✔
Kim *et al.* 1995[168]		✔	
Luker *et al.* 1996[179]			✔
Plumley *et al.* 1995[180]	✔		
Roche *et al.* 1996[181]			✔
Takayasu *et al.* 1995[182]	✔		
Yamashita *et al.* 1996[64]		✔	
Total	2	3	5

TABLE 37 *The three excluded SCT PE diagnostic performance studies*

Reference	≤ 10 patients	Non-human	Duplicate
Blum *et al.* 1994[183]	✔		
van Rossum *et al.* 1996[184]			✔
Woodard *et al.* 1995[173]		✔	
Total	1	1	1

TABLE 38 *The 24 excluded EBCT diagnostic performance studies*

Reference	≤ 10 patients	Non-human	Duplicate	Follow-up < 1 year (asymptomatic subjects)	No CAC score (symptomatic subjects)	Time elapsed > 1 week (reproducibility)
Detrano *et al.* 1995[185]		✔				
Detrano *et al.* 1996[186]			✔			
Fallavollita *et al.* 1996[187]			✔			
Goel *et al.* 1992[188]				✔		
Guerra *et al.* 1993[189]				✔		
Gutfinger *et al.* 1996[174]		✔				
Hoeg *et al.* 1994[190]	✔					
Janowitz *et al.* 1991[191]						✔
Janowitz *et al.* 1993[192]				✔		
Kaufmann *et al.* 1995[193]			✔			
Maher *et al.* 1996[194]				✔		
Mahoney *et al.* 1996[195]				✔		
Mautner GC *et al.* 1994[196]		✔				
Mautner SL *et al.* 1994[197]		✔				
McCollough *et al.* 1995[198]		✔				
Megnien *et al.* 1992[199]				✔		
Megnien *et al.* 1996[200]				✔		
Moshage *et al.* 1995[201]					✔	
Roig *et al.* 1989[202]					✔	
Rumberger *et al.* 1994[203]		✔				
Rumberger *et al.* 1995[204]		✔				
Simon *et al.* 1995[205]				✔		
Simons *et al.* 1992[206]		✔				
Wong *et al.* 1994[207]				✔		
Total	1	8	3	9	2	1

TABLE 39 *The ten EBCT diagnostic performance studies excluded from quantitative meta-analysis*

Reference	Inadequate gold standard	Insufficient raw data	Definition of dichotomy not comparable with those included
Agatston *et al.* 1990[22]	✔		
Agatston *et al.* 1994[102]		✔	
Bielak *et al.* 1994[92]		✔	
Braun *et al.* 1996[94]		✔	
Devries *et al.* 1995[97]			✔
Kajinami *et al.* 1995[70]			✔
Kaufmann *et al.* 1995[103]		✔	
Shields *et al.* 1996[104]	✔		
Tanenbaum *et al.* 1989[100]			✔
Wong *et al.* 1994[105]		✔	
Total	2	5	3

TABLE 40 *The 25 excluded SCT radiation dose studies*

Reference	No comparison with either modality or protocol	No dose measurement or information
Costello and Gaa 1995[128]	✔	
Craven *et al.* 1995[208]		✔
Dula *et al.* 1996[209]	✔	
Engeler *et al.* 1994[210]	✔	
Kalender 1994[2]	✔	
Kasales *et al.* 1995[211]		✔
Luker *et al.* 1993[212]		✔
Moore MM *et al.* 1981[213]	✔	
Moore SC *et al.* 1983[214]	✔	
Nambu *et al.* 1995[215]	✔	
Nishizawa *et al.* 1996[216]	✔	
O'Brien *et al.* 1995[217]		✔
Plumley *et al.* 1995[180]		✔
Reynolds *et al.* 1995[218]	✔	
Robinson *et al.* 1986[219]	✔	
Rubin *et al.* 1993[220]		✔
Ruegsegger *et al.* 1996[221]		✔
Suojanen and Regan 1995[222]		✔
Tsuchiya *et al.* 1994[223]		✔
Vade *et al.* 1996[224]	✔	
van der Bruggen-Bogaarts *et al.* 1996[225]		✔
Vannucchi *et al.* 1996[226]		✔
Villafana 1991[227]	✔	
Vock and Soucek 1993[228]		✔
Wang and Vannier 1994[229]		✔
Total	12	13

TABLE 41 *The two excluded EBCT radiation dose studies*

Reference	No dose measurement or information	Duplicate study
Brody *et al.* 1989[230]	✔	
Zink and McCollough 1994[231]		✔
Total	1	1

Chapter 6

Results of quantitative data synthesis

In this chapter the results of the quantitative synthesis of data, the methodology for which was described in chapter 3, are presented. The results of individual included studies are to be found in chapter 4 and are discussed there and in chapter 8.

SCT – liver lesions

No studies fulfilled the inclusion criteria for a quantitative analysis. A large range of reference standards was used to compare the SCT results, including the biopsy of single lesions, other imaging tests or, in some cases, limited histological correlation.

SCT – PE

Four studies[20,69,87,88] satisfied the quantitative inclusion criteria. They may be divided into those that reported results for the central vessels separately[20,69,87] from results for all vessels (both central and peripheral[87,88]).

Central PE

Two of these studies were by the same research group, but performed on different patients at different times.[20,69] The results of the earlier study[20] were artificially enhanced because inconclusive examinations were excluded from the analysis. The results are given in *Table 42* and an ROC scatter plot is shown in *Figure 7*.

Central and peripheral PE

One of these studies[88] produced two sets of information from two independent observers, which are shown separately in *Table 43*. An ROC scatter plot is shown in *Figure 8*. Because there were so few studies, no attempt was made to synthesise these results quantitatively.

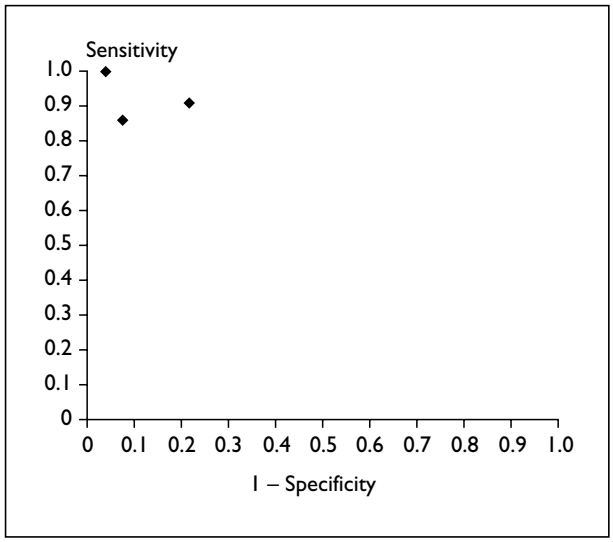


FIGURE 7 *ROC scatter plot: SCT central PE*

EBCT – CAD

To evaluate diagnostic performance, the diagnosis by EBCT is compared with that of the gold standard, angiography. In total, 12 studies performed such a comparison. However, several differences are evident between them, thus preventing direct comparison. The most important is the range of thresholds used for diagnosing the presence or absence of CAD, both for the choice of calcium score cut-off with EBCT and for the severity of stenosis on angiography. For the calcium score, the most common choice was zero for normal, with greater than zero indicating a positive diagnosis of CAD. In the case of angiography, two thresholds were used most frequently:

- zero stenosis for normal against any degree of stenosis for positive CAD
- < 50% diameter stenosis for normal against ≥ 50% diameter stenosis for clinically significant CAD.

TABLE 42 *Results for central PE detection*

Reference	TP	FN	FP	TN	FPR	TPR	OR
Goodman *et al.* 1995[87]	6	1	1	12	0.08	0.86	72.00
Remy-Jardin *et al.* 1996[69]	39	4	7	25	0.22	0.91	34.82
Remy-Jardin *et al.* 1992[20]	18	0	1	23	0.04	1.00	N/S

TABLE 43 *Results for central and peripheral PE detection*

Reference	TP	FN	FP	TN	FPR	TPR	OR
Goodman *et al.* 1995[87]	7	4	1	8	0.11	0.64	14.00
van Rossum *et al.* 1996[88]	10	5	0	41	0.00	0.67	N/S
van Rossum *et al.* 1996[88]	12	3	0	41	0.00	0.80	N/S

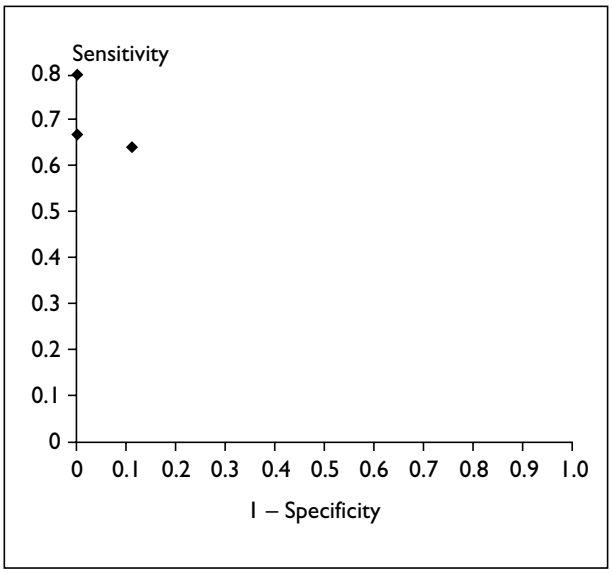


FIGURE 8 *ROC scatter plot: SCT central and peripheral PE*

Three studies used the first criterion and seven studies used the second. The following analysis was performed on the larger data set of seven studies.[91,93,95,96,98,99,101] Amongst these, although they used the same diagnostic criterion, many other differences were present. These included patient characteristics, the number of slices, and the number of pixels used to identify the presence of calcium. When combining or comparing these studies, these differences must be taken into account. An assessment of the influence of study characteristics appears in chapter 7, using the methodology described in chapter 3. The small number of studies available means that confidence in the results is low.

The raw data for these seven studies are shown in *Table 44,* together with the calculations for TPR, FPR and OR. The corresponding values, FPR2, TPR2 and OR2, are those obtained after the addition of 0.5 to each of the cells of the 2×2 contingency table.

The results given in *Table 44* are shown on an ROC plot (*Figure 9*). It can be seen that several studies have a specificity below 50%. Moses *et al.*[45] has suggested including only studies above a clinically appropriate threshold of 50% for sensitivity and specificity, but this was not implemented in this analysis because the number of studies was already so small.

The influence of excluding one particular study was assessed. The article by Budoff *et al.*[96] reported a multicentre study and incorporated results from some of the centres that had been published individually in the other articles. In effect, the results were being counted twice. The data were plotted twice after logistic transformation (*Figure 10*), and an SROC calculated using both methods of fit: EWLS and RR.

Table 45 shows the linear fit parameters associated with the SROC curve. For the EWLS fit, the influence of excluding the multicentre study has no significant effect on either the gradient or the intercept. For the RR fit, the gradient remains the same both with and without the multicentre study, but the intercept differs. This difference can be attributed to the inherent insensitivity of the RR fit. In the RR fit, the intercept of the line is determined by placing the line midway between points, and so the standard error (Se) of such a technique is difficult to estimate. This is not the case for the EWLS method. The results from the EWLS fit were therefore used in the remainder of the analysis. The multicentre study was included in further analysis because it also included original data that were not supplied by any other study.

Using equation 2 (p. 21) and the gradient and intercept from the EWLS fit, the data were transformed back to ROC space (*Figure 11*). To provide a statistic summarising this ROC curve, the TPR was read from the mean FPR (0.51): sensitivity = 95% at specificity 49%.

TABLE 44 *Statistical results of EBCT studies*

Reference	TP	FN	FP	TN	FPR	TPR	FPR2	TPR2	OR	OR2
Barbir *et al.* 1994[91]	15	3	31	53	0.37	0.83	0.37	0.82	8.55	7.52
Bormann *et al.* 1992[93]	15	1	25	9	0.74	0.94	0.73	0.91	5.40	3.85
Breen *et al.* 1992[95]	47	0	28	25	0.53	1.00	0.53	0.99	N/S	85.00
Budoff *et al.* 1996[96]	404	23	159	124	0.56	0.95	0.56	0.95	13.70	13.44
Fallavollita *et al.* 1994[98]	50	9	26	21	0.55	0.85	0.56	0.84	4.49	4.31
Rumberger *et al.* 1995[99]	64	1	16	30	0.35	0.99	0.35	0.98	120.00	79.49
Yaghoubi *et al.* 1995[101]	32	1	15	19	0.44	0.97	0.44	0.96	40.53	27.26

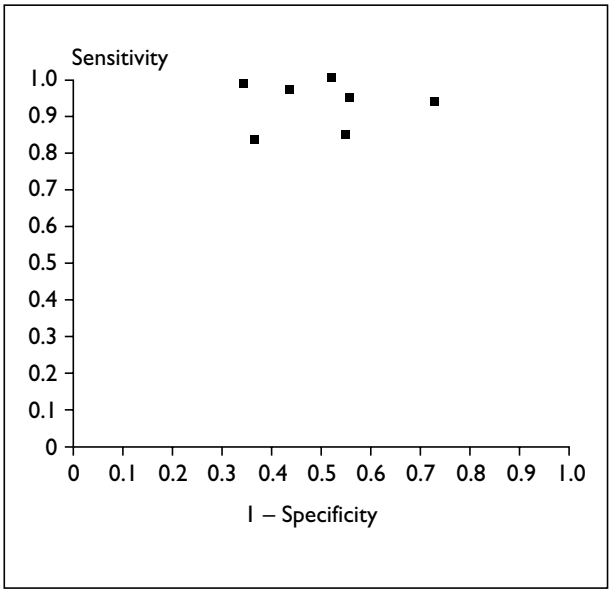


FIGURE 9 *ROC scatter plot: EBCT for symptomatic CAD*

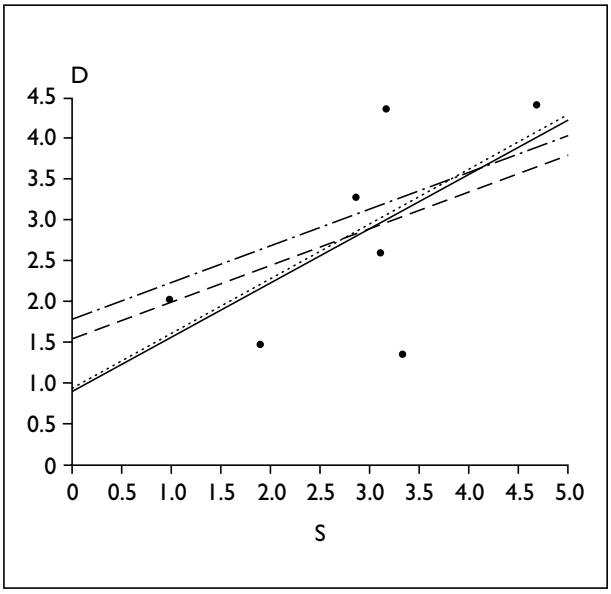


FIGURE 10 *SROC curve (logistic transformation): EBCT for symptomatic CAD (——, EWLS7; – – –, RR7; ·····, EWLS6; – · – ·, RR6)*

TABLE 45 *EBCT using calcium score for diagnosis of symptomatic CAD: results of linear fit of logistic SROC plot (n = 7 with the multicentre study; n = 6 without the multicentre study; see chapter 3 for method)*

Method	B	A	Se(A)
EWLS (*n* = 7)	0.67	0.88	1.19
EWLS (*n* = 6)	0.68	0.91	1.32
RR (*n* = 7)	0.45	1.55	–
RR (*n* = 6)	0.45	1.80	–

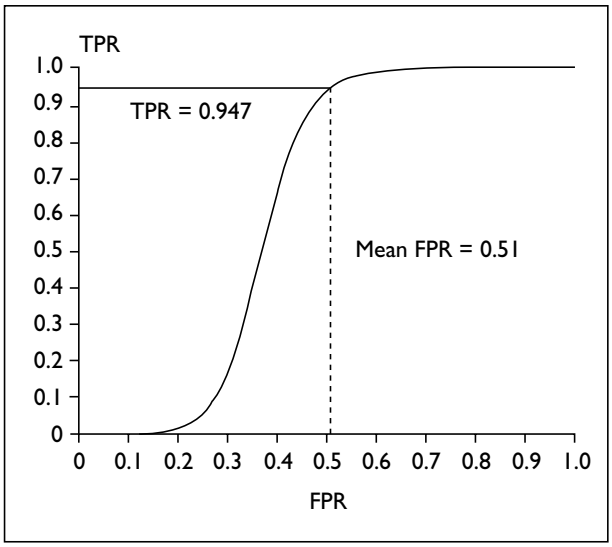


FIGURE 11 *SROC curve: EBCT for symptomatic CAD using the EWLS method*

Chapter 7

Analysis of the robustness of the results

The quantitative results presented in chapter 6 represent the most reliable available evidence about diagnostic performance. However, the studies included in the analysis were not all identical (in design, equipment used, or patient selection), nor were they free from threats to validity in the form of biases. In this chapter the effect of these differences on the results of the quantitative analysis is investigated.

A notable feature is the poor quality of the original studies, in terms of study design and, in particular, the completeness of the subsequent published reports. Using the checklists described in chapter 3, each study was assessed for the risk of bias and the qualitative factors were noted. Rather than exclude studies on the basis of this information, further analysis was undertaken to determine if the results were related to these factors, or to the risk of one or more biases. The methodology was described in chapter 3 and expands on the SROC analysis technique.

The analysis could be applied only to the topic of the diagnosis of CAD in symptomatic patients by using EBCT because this was the only area where there were enough suitable and similar studies to perform quantitative analysis. Two potential biases and three factors were evaluated (*Table 46*). Note that the 'blinding biases' category covers all four biases listed under the heading 'Independence of interpretation' in *Table 6*. The mention of blinding of any information for any of these four biases was interpreted as indicating no risk of bias for the regression analysis. The number of pixels is a criterion used as part of the calcium scoring method to help to differentiate small calcified lesions from image noise. The number of slices refers to the coverage of the coronary arteries with the CT acquisition. If more slices are used, there may be greater coverage to ensure the depiction of peripheral arteries, or the slices may be thinner than those used in other protocols. In the studies included, the slice thickness was not changed, nor were overlapping slices used, so the result of using more slices was increased coverage.

TABLE 46 *Biases and factors included in the analysis showing how values were assigned to one of two categories: this division is described in the text as the dichotomy; 'no', 'yes' and '?' are the alternatives available in the checklist*

Factor or bias	1	0
Disease progression bias	No	Yes or ?
Blinding biases	No	Yes or ?
No. patients	> 100	< 100
No. pixels	2	> 2
No. slices	At least some > 20	≤ 20

Other patient characteristics such as age and sex were not included because these factors were evaluated as part of the subgroup analyses performed within the studies. The full list of biases and factors could not be evaluated because some did not vary between studies. For example, all studies supplied insufficient information adequately to assess referral bias, while all were free from incorporation bias. For the multiple regression analysis it was necessary to divide the values for each bias or factor into two categories representing the binary alternatives 0 or 1 (*Table 46*). All variables were placed in a multiple regression with a 5% ($p < 0.05$) level of significance for inclusion. No variables were found to be significant. Regression was then performed individually on each variable using the same level of significance ($p < 0.05$).

When testing multiple variables, if n tests are performed, it is recommended[232] that the desired significance level is divided by n to arrive at the uncorrected probability that should be used to determine statistical significance. This is sometimes known as the Bonferroni correction. In this case, a value of $p < 0.01$ would be required. No variables met this new stricter significance criterion. The 5% level of significance was used and the variables that individually met a 5% significance level were analysed. These results must therefore be interpreted cautiously because the correction has not been applied.

The data for the seven studies concerning the EBCT diagnosis of symptomatic CAD are shown in *Table 47*.[91,93,95,96,98,99,101] Only the number of slices was significant at the 5% level ($p = 0.044$). In the SROC plot for this regression analysis (*Figure 12*), the studies are separated into two

TABLE 47 *EBCT diagnosis of symptomatic CAD: factors and biases for the studies in the meta-analysis*

Reference	Disease progression	Review bias	No. patients	No. pixels	No. slices
Barbir *et al.* 1994[91]	Yes	No	> 100	2	40
Bormann *et al.* 1992[93]	No	?	< 100	2	20
Breen *et al.* 1992[95]	No	No	> 100	2	40
Budoff *et al.* 1996[96]	No	Yes	> 100	2 or 4	> 20
Fallavollita *et al.* 1994[98]	No	No	> 100	4	20
Rumberger *et al.* 1995[99]	No	?	> 100	2	40
Yaghoubi *et al.* 1995[101]	Yes	?	< 100	2	20–40

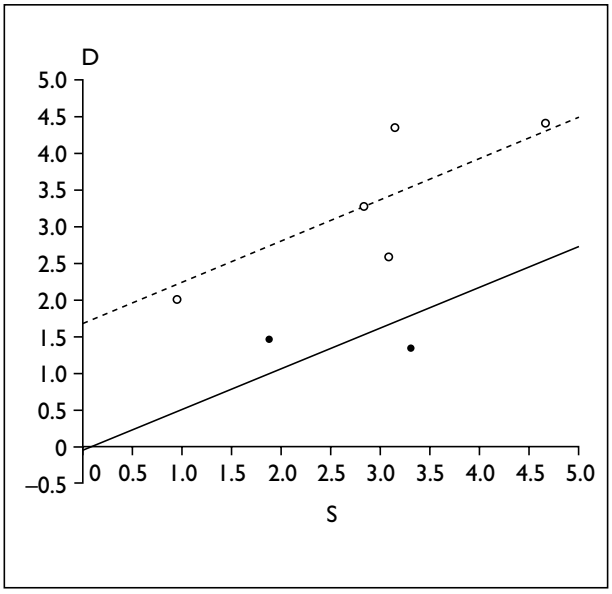


FIGURE 12 *SROC curves (logistic transformation) for studies using different numbers of slices: EBCT for symptomatic CAD (●, 20 slices; ○, 20–40 slices; ——, 20 slices; ·····, 20–40 slices)*

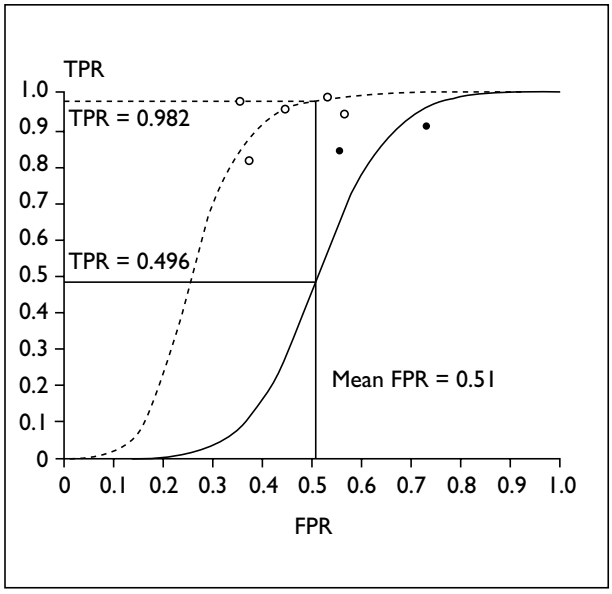


FIGURE 13 *SROC curves showing the effect of the number of slices on diagnostic performance: EBCT for symptomatic CAD (●, 20 slices; ○, 20–40 slices; ——, 20 slices; ·····, 20–40 slices)*

sets, one for the results with 20 slices only and one for those with 20–40 slices. In this method of analysis, the gradient of the line for each set is the same and the intercepts differ (*Table 48*). The intercept, A, is the estimated ln(OR) when sensitivity equals specificity (S = 0). The separate sets were transformed back to conventional ROC space (*Figure 13*) and summary estimates of TPR from the mean FPR made. A clearer representation of the effect of the number of slices on the accuracy of diagnosing CAD can be seen: increasing the number of slices improves diagnostic performance.

TABLE 48 *Effect of the number of slices acquired on the diagnostic performance of EBCT calcium score for CAD (see chapter 3 for method)*

Method	B	A	Se(A)	TPR (mean FPR = 0.51)
Total (*n* = 7)	0.67	0.88	1.19	0.947
20 slices (*n* = 2)	0.57	–0.07	0.82	0.496
20–40 slices (*n* = 5)	0.57	1.68	0.80	0.982

Chapter 8
Discussion

In this chapter, the review methodology is discussed, with particular reference to potential biases in the approach. The results of the review reported in chapters 4 and 6, and analysed for robustness in chapter 7, are discussed with reference to the questions posed in chapter 2. This chapter concludes with an overview of the development of the literature in this field.

Review methodology

Search strategy

The search strategy was designed to have a high recall and therefore low precision. No limitations on publication type, such as the search for randomised controlled trials,[28] were applied because the available data were limited. In addition to MeSH classifications for CT, alternative keyword searching was used to identify latest generation CT studies. A possible limitation of using keyword searching for this purpose is that, because this evolving technology is based on existing technology, not all studies may specify the new terminology in the title or abstract and may therefore be classified as the old technology. This hypothesis is more applicable to SCT because it is widely available and has become generally accepted as the norm.

Because of the diversity of CT technology and the range of terminology that is used medically, no attempt was made to search electronically for the specific topics of liver lesions, PE or CAD. Instead, the abstracts of those studies retrieved from the keyword search for latest generation CT were systematically read and classified. This method is labour intensive and requires the full publication to be acquired if no abstract is provided. However, the system used has the following advantages.

- It does not rely on the accuracy of MeSH terms.
- It is still efficient if no subject classifications are available.
- It allows the use of as many individually designed classifications as required.
- Once the process is complete, if a database of classifications is maintained, a record of both included and excluded data, together with reasons, can easily be accessed and searched.
- No references are discarded, therefore an extension of the coverage of the review will not require any further searches to be performed.

When searching for evidence on the higher levels of the hierarchical framework, the same process, using suitable keywords, was applied. Although the searches were more general than at the diagnostic performance level because no specific clinical application area was considered, more MeSH terms, such as 'outcome and process assessment (health care)', were used to increase the precision of the search for studies at the required levels of the hierarchical framework.

In contrast to the findings of other workers, all but one of the studies identified were in MEDLINE. For example, Dickersin *et al.*[233] reported the sensitivity of MEDLINE to be as low as 51% for randomised controlled trials in ophthalmology. Further, in our review of the health economics of latest generation CT, MEDLINE supplied all the information. On re-analysing the occurrence of the keywords and MeSH terms used in the search, it was found that the MeSH term 'economics' was sufficient to retrieve the two studies we identified as the best available evidence. Indeed, this finding was repeated in another review.[31]

The one study not contained in MEDLINE was found from handsearching the cited reference lists from retrieved articles. No other search method retrieved any new studies. Only the reference lists from studies that were retrieved, and therefore of potential value, were searched. Review articles, which are rich sources of references, were not systematically searched because a large number exist and, with limitations on time, it was believed that searching other sources would be more profitable. Although a rigorous comparison has not been made in this review, the finding suggests that time and effort may be better spent conducting an exhaustive search of all articles' reference lists, particularly review articles, than in searching EMBASE and other sources. The cost of retrieving more articles to search their reference lists should be weighed against the cost of searching databases. Retrieval costs would be reduced if more databases provided lists of references on-line, similar to the service provided by BIDS-ISI.

A decision was made not to write to authors of abstracts to determine whether their work had subsequently been published because we had experienced a very low response rate to this tactic when conducting a previous review.[31] In addition to ensuring the completeness of the review, such mailing would also allow an estimation of the degree of publication bias. In the previous review, however, the topic was such that equivocal or negative findings were unlikely, and it could be hypothesised that no articles would fall into a category that would suffer from rejection because of the findings. We are aware that in the current review, especially in the SCT area, the potential for publication bias must exist. For example, an article showing that the performance of SCT was about the same as conventional CT would be less likely to be published than one showing markedly better performance. A full examination of potential publication bias was beyond the scope of this review.

Inclusion criteria

The quality of a review depends on the quality of the primary studies. The inclusion criteria must be based on the level of evidence available; for example, setting an inclusion criterion of a randomised controlled trial design for this review would not have been profitable. The type of study available led to inclusion criteria that were primarily not concerned with the study design. The criteria chosen were sufficient to allow an adequate comparison of studies in terms of the presentation of results as well as the field of the study. For the studies assessing the higher levels of the evaluative framework, the criteria were less stringent because of the lack of data. The criteria were set to include studies that supplied any information, irrespective of the study design or quality.

The decision to exclude non-English language studies is a potentially biasing influence in the design of our review. This decision was made for pragmatic reasons because we had difficulty in identifying technical translators willing to undertake this work. Full translations would have been essential because the abstract rarely supplied sufficient information for a full assessment. One possibility was to use non-specialist overseas students to read the untranslated articles and extract the information required. However, because of our experience with the difficulty of assessing English-language articles, this route was not pursued. In this topic area the range of languages was relatively small. For example, our database of articles on EBCT in CAD comprises 13 that are non-English language: six German, five French and two Japanese. These represent articles that have, on the basis of their abstract, satisfied the preliminary inclusion criteria and also the topic-specific criteria. The next stage would be to obtain copies of the full articles, and translate and re-check against the criteria.

To test the hypothesis that non-English language articles might be published twice, in English and in another language, this list was compared with our included and excluded reference list. The hypothesis was not substantiated, because only one group of authors appears in both lists. This finding adds to the view that the omission of the non-English language papers may have a significant effect. A Health Technology Assessment Programme research topic priority was identified in 1996: 'The inclusion of non-English language trials in systematic reviews', but, at the time of writing, work had not commenced.

Assessment of the relevance and validity of primary studies

In the analysis presented in chapter 7 it was stated that there was no statistically significant relation between the bias risk and the study result. This finding is not conclusive because data were missing from the regression analysis. In common with other authors,[43] we chose to describe a bias risk as present if the information required to determine its presence was not given in the article. Omissions in the descriptions of the primary studies were so widespread that it was not practical to contact the authors to confirm their exact methodology. There are two possible outcomes.

- A proportion of studies in which there was no risk of bias have been misclassified. This could prevent any relationship being significant at the specified level.
- The bias risk was present in almost all of the studies, meaning that there were insufficient studies without the bias risk for discrimination in the analysis.

The investigation of EBCT for its accuracy in detecting symptomatic angiographic CAD was the only topic in the review where sufficient similar studies were available for a quantitative data synthesis to be performed. It was noted in chapter 7 that the full list of biases and factors could not be evaluated because some did not vary between studies. This statement merits further discussion. The completed bias checklist results are given in appendix 4 (*Table 52*),[91,93,95,96,98,99,101] and the questions that the reviewers answered are in the bias checklist for diagnostic performance in appendix 3. It must be emphasised that similarity between

studies is not necessarily a fault, because the authors may all have taken measures to minimise the risk of bias in their results. In this class, of the seven articles, six or more gave sufficient information to determine that measures had been taken to minimise the risk of bias in the following categories: verification, work-up and incorporation biases; withdrawal bias; and comparator review bias. Thus, these are areas in which study design and reporting were adequate. However, of the seven articles, six or more did not supply information for assessment, or had not minimised the risk of bias in the following categories: referral bias, patient filtering bias, intra- and interobserver variability, and clinical review bias.

It could be argued that it is not surprising that different studies produced different results, if they used, for example, different equipment, or investigated completely different sets of patients (these are examples of the variables we have called **factors**). We chose to include all studies in the quantitative analysis, because of the wide range of studies described in the literature and the likelihood of being unable to perform any synthesis of results without the regression approach. This had the potential additional advantage of giving quantitative evidence of the impact of any differences.

Data extraction

Our review methodology was designed to minimise bias by using a multidisciplinary panel drawn from more than one centre. After an initial phase to check inter-reviewer reproducibility, the data extraction was performed primarily by a single reviewer, who consulted other panel members if in doubt. It is possible that this strategy may have introduced a degree of bias into our review process, but it is not believed to be large because, in the event, the more difficult and potentially subjective decisions (such as bias risk) were not used as inclusion criteria.

Data synthesis

The number of studies suitable for quantitative data synthesis was small. Therefore the following conclusions on the applicability and validity of the methodology are only suggestive and require further research with larger data sets to confirm their importance.

The first point worth noting is the shape of the SROC curve in *Figure 11*. The right-shifted shape observed is characteristic of a screening test, where high sensitivity is required even at the expense of test specificity, although the results shown are for a test performed on a symptomatic population. The choice of the summary statistic for the SROC curve is not straightforward. Quoting a statistic such as Q* may not be the appropriate choice because this balances sensitivity and specificity. All operating points along the curve are valid and the summary statistic should reflect the best application of the test. Therefore, a better summary statistic may be the sensitivity read from the mean specificity, which will reflect the threshold range chosen by the studies.

The suggestion by Moses *et al.*[45] of choosing a 'clinically relevant' range for both sensitivity and specificity was not used in the analysis. Owing to the small number of studies available, any further exclusions were undesirable; secondly, the choice of the relevant range is subjective and not easily defined for the clinical application.

Although the SROC methodology has been applied here, it is doubtful whether it is entirely appropriate in this case. The method involves the addition of 0.5 to all the cells of a 2×2 matrix in order to avoid zero cells. Such zero cells occur when sensitivity or specificity equal 100% or 0%, which would prevent the use of equation 1 (p. 21), because of a zero denominator. Although the addition of 0.5 allows the analysis to be performed, it does affect the sensitivity and specificity. An extreme example is if the results from several studies, all with 100% sensitivity, were combined. The combined sensitivity returned by the analysis would be less than 100% because of the addition of 0.5. The amount by which the sensitivity or specificity is affected by the addition of 0.5 is dependent upon both the number of patients and the value of sensitivity or specificity. In the case where sensitivity and specificity are approximately equal, both parameters will be equally affected by adding 0.5. However, as in this case, when one is high (sensitivity close to 100%) while the other is intermediate (specificity 50%), the addition of 0.5 has an unbalancing effect. This imbalance may possibly distort the shape of the SROC curve.

Health economics

The immediate conclusion from the review of the literature is that there are no comprehensive economic studies of SCT or EBCT from which conclusions about cost-effectiveness can be drawn. However, some useful information on the direct costs of the procedure can be located and, as discussed later in this chapter, some data on diagnostic and therapeutic impact can be found. Data on patient outcomes that are suitable for

more sophisticated types of economic analysis, such as cost–utility analysis, are not available. Given the disparate sources of this information, the only feasible approach to the economic evaluation of SCT and EBCT at present is by decision modelling, as illustrated by van Erkel and colleagues.[19]

This approach is familiar to economists and is being used increasingly in the clinical field. The principal advantage of modelling is that the analysis can be designed to address a specific question and can draw the best available data from multiple sources. The main disadvantage is the need to make assumptions when suitable data do not exist. Although the effect of assumptions can be tested through sensitivity analysis, the danger of bias being introduced is real. Good reviews of the use of modelling in economic evaluation in health care can be found in Sheldon[234] and Buxton *et al.*[235] The general agreement is that modelling is useful in generating and selecting hypotheses prior to the design of clinical trials and in extending the application of clinical trial data to different patient groups, care settings, and time periods. This last point is particularly important in the evaluation of diagnostic techniques because the time between investigation and final patient outcome is often quite long and few trials follow up for that length of time.

SCT

One study of the cost-effectiveness of SCT versus CTA in the diagnosis of PE[19] was well conducted and clearly presented. A decision model was used, drawing diagnostic accuracy and prognostic data from the literature and costs from a specific hospital. Outcomes were measured in lives saved rather than QALYs because of the lack of available data on QoL.

The only study of malignant liver lesions with significant economic content was that of Semelka *et al.*,[50] which compared CTAP with MRI. The costs of the investigations were estimated from hospital billing data. Although fewer than half the patients had surgical validation of diagnostic findings, CTAP was judged more sensitive but less specific than MRI. The referring surgeons judged that CTAP did not alter management as determined by MRI findings. Because CTAP was also more costly, it was not recommended as a replacement for MRI. Given the small number of patients ($n = 26$), the presence of verification bias, the observational design of the therapeutic impact assessment, and the restricted nature of the costing, it would be unwise to base firm conclusions on these results.

The other two studies were a comparison of investigation strategies to evaluate living renal donor candidates[49] and a comparison of SCT and radiography in screening for lung cancer.[48] Lindgren *et al.*'s study[49] identified potential cost savings for the approaches including SCT, but used charge data and a limited range of costs. Kaneko *et al.*'s study[48] used patient payments as a proxy for procedure costs and did not follow up the impact on management and outcome. Neither of these studies is a suitable basis for informed decision making.

EBCT

The single economic study of EBCT consists of a short note describing the set-up and operating costs of a new EBCT unit in France.[51] The EBCT cost data could be useful for other studies, but equivalent comparative data are not given for conventional CT, making the average cost difference quoted difficult to interpret.

Patient outcome

SCT

There is a dearth of evidence in this area. No studies at all were included relating to patient outcome and SCT.

EBCT

Three studies were included in the EBCT topic area. Two of these concerned the detection of CAC. In one,[52] the results suggested that EBCT imaging at the time of PTCA may predict the likelihood of re-stenosis after the procedure. In the other,[54] the emphasis was on behavioural changes caused by the results of EBCT examinations on asymptomatic but self-referred individuals, where a significant link was found. In the third article,[53] EBCT was used in the staging work-up of patients with hypopharyngeal squamous cell carcinoma. The results indicated that EBCT did contribute to better patient outcome.

It is interesting to note that none of these applications of the technology is in mainstream use. Particularly noticeable is the lack of outcome studies relating to the use of EBCT for screening the asymptomatic population for CAD.

Therapeutic impact

SCT

These three studies covered different clinical applications. The first[55] is relatively anecdotal,

describing changes in management in three of 11 patients in whom SCT of the tracheobronchial tree was performed. A larger (n = 1879), retrospective study[56] looked for incidental findings of PE. Patient management was changed in 11 instances. The third study[50] investigated the technology applied to the management of malignant liver lesions. The CT findings were compared with those from MRI and did not change management in any patient.

The topic area of liver lesions was one on which this review focused as a specific clinical application. Seven studies related to diagnostic performance were included in the review, but none of those attempted any investigation of therapeutic impact. This is a typical finding in medical imaging. One recommendation of this review is for educational initiatives to inform investigators of simple enhancements to study protocols to allow the assessment of therapeutic impact.

EBCT

Again, three studies were found and, as for the patient outcome level, none was on a mainstream EBCT topic. The overall conclusion was positive if relatively anecdotal. EBCT was used in the management of bone marrow transplant patients in whom chest radiography had been inconclusive; a statistically significant benefit was found.[57] In the evaluation of intracardiac masses[58] there was evidence that EBCT made a significant contribution to patient management decisions, but the study did not clearly identify changes in management resulting from the use of this modality. A small study[59] involving parotid masses found changed management in two of 13 patients.

Diagnostic impact

No studies specifically designed to evaluate diagnostic impact were found. As a secondary standard, those comparing modalities with an independent reference test were reviewed.

SCT

For SCT of liver lesions, MRI was the comparative modality most used, followed by conventional CT and US. For MRI, spin echo, dynamic and RARE techniques were used. A wide variety of procedures and protocols were presented in a small number of articles.[60–64] The performance of MRI in comparison with SCT was reported as being better, equal or worse. Two of the MRI studies also compared with US. One[60] suggested US to be an inferior test to MRI and SCT, and the other[62] suggested that MRI and US were equivalent. A failing in the descriptions of these studies was the lack of information given on the US methodology.

In comparison with conventional CT, two studies[65,68] suggested that image enhancement was superior for SCT, particularly for the last slice in the dynamic series. The other two studies[66,67] investigated performance in detecting liver lesions, with both suggesting that SCT was better for lesions smaller than 20 mm. Again, differences between studies, the small number of studies, and the sizes of patient groups limit the conclusions that may be drawn.

Only one study[69] was included in the PE topic area; most comparative studies for SCT in PE lacked a comparison of both tests with an independent reference. This study provided comparative information with the most popular test for PE (i.e. VP), using pulmonary angiography as a reference standard. The conclusion was that SCT was better than VP, but this was based on a small subgroup of patients. The choice of gold standard might affect such comparisons: one could predict better agreement for SCT with pulmonary angiography because these two techniques both image filling defects in the arterial lumen, while VP allows visualisation of the effect of defects on pulmonary function.

EBCT

For EBCT, fewer comparative modalities were presented. This may be because EBCT is the only modality that can quantify calcium in terms of a score as an indirect diagnosis of CAD. Other modalities primarily measure the degree of stenosis. One study[22] presented a subgroup of patients who also underwent fluoroscopy. It was suggested that EBCT was superior, but this was only the case for sensitivity. The specificity was better for fluoroscopy. The other study[70] compared EBCT, ECG and exercise thallium stress tests, both independently and in combination. In most patient groups the specificity for EBCT was higher than for the other tests, but with lower sensitivity. We found no evidence directly comparing the predictive performance of modalities.

On commencing this review, it was expected that evidence would be lacking, particularly at the higher levels of the evaluative framework. It was unexpected, however, that so little evidence was available about diagnostic impact, as it would seem to be a natural next step to determine both how the technology was affecting diagnosis and its relationship to other investigations.

Diagnostic performance

One of the factors associated with the reporting of SCT is the manner in which it is performed. Different results would be expected when studies are read on the console, where the reader has control of windowing, than when pre-windowed films are used. Limited information about the viewing methodology was given in the publications.

SCT – liver lesions

Comparison of HAP with PVP

Our aim here was to determine evidence that additional lesions can be seen when using the HAP compared with the PVP. All four included studies[71–74] reported findings along these lines, with percentage increases in terms of lesions of between 8% and 11%. A variety of lesion types were included.

HAP versus reference (for HCC only)

These four studies[61,62,71,75] showed that performance depends on lesion size; it is reduced below 1 cm. Lesion detection results are quoted in percentage terms, giving the number detected out of the number seen when using the reference techniques. For all sizes of lesions, results between 52% and 88% for a full protocol including the HAP, and between 47% and 84% for using the HAP alone, were reported. A range of reference standards was employed within and between studies, as well as a range of acquisition protocols, so the wide ranges are not surprising.

Studies designed to compare PVP protocols

One difficulty with the studies in this category is the use of maximal parenchymal enhancement as a measure of the conspicuity of hypovascular lesions. This encourages the selection of high intravenous contrast dose protocols, when it may in fact still be possible to detect a lesion at lower contrast doses. It may be seen in *Table 20* that, in all studies, the highest contrast doses gave most enhancement. Agreement between studies was less apparent for the other parameters investigated. Two studies[80,82] found that a faster injection rate increased enhancement; one[79] found this with a slower rate. Two studies[77,80] indicated a preference for a uniphasic (constant rate) injection rate, while a third[79] advocated a biphasic rate (relatively rapid at first, then slower). Four[78–81] of five studies found that the longer delay time used was better; the fifth[68] gave non-significant results.

A major difficulty in comparing the results from the different studies was the variety of uncontrolled intrinsic factors involved. Dodd and Baron[76] listed these as including: the patient's weight, cardiac function, state of hydration, renal function, time since last meal, primary liver disease, and compromise of the hepatic vascular supply. Patient weight was not detailed in four of the seven studies (*Table 16*).

Although the studies on automated scan technology differed in approach, they all found that the automated technique was better than using a fixed scan delay. It is likely that this method will become standard practice in future, with the exact protocol being dependent on the manufacturer. This point is addressed further in chapter 9.

SCT – PE

Our aim was to determine if more information was available than at the time of van Erkel *et al.*'s study.[19] Of the five studies listed by these authors, one was excluded from our review because fewer than ten patients were involved,[183] and a second because it was a German language article.[236] We included two from van Erkel *et al.*'s list,[20,87] one updated version[88] and a single additional article.[69] Thus, there has been little change in the published literature since that review, and our inclusion criteria appear to have been broadly similar.

The first study, from 1992,[20] gave enhanced results owing to the exclusion from the analysis of those studies with inconclusive CT findings. Two studies[69,87] gave results for central vessels only, in which sensitivity ranged from 86% to 91% and specificity from 78% to 92%. All the studies emphasised that SCT may fail to depict small peripheral emboli, and this problem accounts for the lower values for sensitivity and the higher ones for specificity if results are calculated for all vessels:[87,88] 63–80% and 89–100%. Although there is clearly still some disagreement about absolute values, and this may be caused by differences between the studies, the future use of SCT in this clinical application depends upon the determination of the clinical importance of peripheral or subsegmental PE. If such clots are found to be of little relevance, then SCT shows promise for a role in the diagnosis of acute PE. If they prove of relevance, then a negative SCT scan should always be followed up with pulmonary angiography, which is more sensitive for such vessel occlusion.

EBCT – CAD

Prediction of asymptomatic CAD

In spite of the perception that this is the major sales application for EBCT, we found only two studies[89,90] to include in the review (*Table 23*).

Ten further studies were excluded because their follow-up period was less than 1 year. It must be hoped that, as time passes, the number of publications documenting a reasonable period of follow-up will increase. Of the included studies, a large one,[89] with 1173 patients, found no correlation between CAC score and subsequent cardiac events. The other[90] was an earlier study that did not use the calcium score methodology. It claimed a 95% agreement between SCT findings and unspecified clinical and laboratory results; this agreement was present at the time of the examination and after 5 years. No information was given about subsequent cardiac events.

Our decision to exclude studies with less than 1 year of follow-up had a major effect on the number of studies included. The relaxation of this inclusion criterion, to include studies that attempted to correlate calcium score with parameters that were measurable at the time of the EBCT study, and not with subsequent cardiac events, is a possibility. Panel discussions, however, led us to the conclusion that such an approach does not really answer the question about prediction, for which a follow-up period is essential.

Accuracy for symptomatic CAD

Thirteen studies were included (*Table 24*).

The choice of EBCT dichotomy has a marked effect on the results of this test. Nine of the included studies classed results as positive if the CAC score was greater than zero. This is a very low threshold and leads to the right-shifted ROC curve seen in *Figure 11* from the quantitative synthesis of data. This is a surprising choice of threshold for a test to be performed on symptomatic patients because it leads to a sensitivity to specificity relation more suited to screening for the presence of disease. If the EBCT is applied instead to rule out the presence of disease, the choice is of less concern.

Five studies[70,92,98,101,105] did look at using a range of thresholds for the positive/negative dichotomy. Bielak *et al.*[92] suggested that a 2 mm^2 area threshold should be used, which meant that, although the sensitivity was not as high as in some reports, neither was the specificity as low. They found that smaller hyper-attenuating foci did not repeat more than 50% of the time on a second examination. They pointed out that, although such noise could be reduced by increasing exposure time, this would increase the partial volume and motion artefact and wipe out any benefit. Fallavollita *et al.*[98] found a similar effect: when they increased the threshold for defining a positive EBCT scan from 1 mm^2 to 2 mm^2, the specificity was increased from 36% to 45%, and the sensitivity fell from 88% to 85%. Kajinami *et al.*[70] reported only the results from using an optimised threshold value determined from ROC curve analysis (they assumed that the negative impact of false-positive results was equal to the negative impact of false-negative results). Their results showed much higher specificity than reported by others (86% over all patients for a sensitivity of 77%) and it must be noted that they were also the only investigators to define angiographic disease densitometrically. Wong *et al.*[105] investigated a very large range of score cut-offs. They showed an increase in specificity from 43% to 71% when the score cut-off was increased to 50, with a decrease in sensitivity from 92% to 78%. Their comparison was not with the angiographic findings alone. Yaghoubi *et al.*[101] also included a comparison using an alternative scoring system; changing the threshold definition from one to eight pixels improved specificity from 59% to 62%, which was not significant.

Although re-analysis of the data in studies in which a low-threshold calcium score has been used to define the CAD would be interesting, the raw data are not available in the published articles. Budoff *et al.*[96] addressed this point in their multicentre study. They found that there was no statistical difference between the sensitivities and specificities reported by centres using different threshold areas.

A further factor is the possibility of the underdiagnosis of CAD by angiography if stenosis is used as the only measure and irregularity is not included. The was specifically suggested by Breen *et al.*,[95] and the investigation by Devries *et al.*,[97] who showed an increase in specificity from 41% to 55% if the identification of any lumen irregularity, not just stenosis, was used to define the presence of disease. However, this was accompanied by a reduction in sensitivity from 97% to 88% and, interestingly, a reduction in the NPV from 94% to 76%.

Six articles concluded that the most promising role for the modality was in ruling out disease, based on high values of the NPV. Only one study[98] indicated that even the NPV was too low, and these patients were in a younger age group. Such conclusions need to be applied with care, given the reported findings[94,103,105] that show how other factors such as age, sex and race can affect the score. Indeed, these relations with coronary calcium were reported well before the advent of EBCT.[237] Braun *et al.*[94] showed a strong positive

relation between age and mean total CAC for three patient groups, regardless of the presence of angiographic CAD. Wong *et al.*[105] demonstrated differences between patient groups defined by age and sex, with less CAC in women and the young. The article by Kaufmann *et al.*[103] plotted percentiles of CAC scores by age and sex, with the aim of determining age- and sex-specific thresholds. Further work along these lines will be essential, together with follow-up studies to include outcome, for EBCT to be properly utilised.

In this review, we have considered the diagnostic performance of the EBCT calcium score for detecting angiographic CAD. It is clear that this does not address adequately the broader issues associated with the use of the technique, particularly the clinical significance of the calcium score. As is the case for studies that aim to predict asymptomatic CAD, a better test would be the occurrence of cardiac events in long-term follow up.

Reproducibility for disease progression

The reproducibility of the measurement can be compromised by the partial volume effect, where small lesions may be imaged, or not, depending upon their precise location within the slice. The presence of cardiac and respiratory motion adds to this uncertainty, in spite of gating and breath-hold acquisitions. Such problems will be worse in patient groups in whom there is less calcium.

Although a variety of statistical parameters and tests were used in the included articles to evaluate their reproducibility, the overall conclusion is that intra-observer reproducibility and interobserver reproducibility for measurements on the same data acquisition are acceptable. However, the five inter-examination results, for examinations repeated within 24 hours, are not encouraging. The authors of two articles felt able to recommend the technique for longitudinal studies. In one,[110] they used a completely different method of analysis, which makes it incomparable with the others. The second[107] investigated methods of improving reproducibility. Most notable was their use of a limited number of slices and the introduction of a threshold value for scoring, thus excluding the troublesome low values. The consensus of the studies that included low calcium scores[106,108,112] was that EBCT was satisfactorily repeatable only for population studies, and that a single examination was inadequate. This is an important finding. A re-analysis of the raw data would be interesting, using the same statistical technique on all sets of results.

Radiation dose

SCT

A large number of studies were excluded from the review because they did not present comparative results from another technique or modality. Such exclusions were necessary because of the wide variety of experimental techniques used and measurements made, which meant that simple comparisons of results would be invalid. Six studies compared the SCT dose for a pitch of 1 with that for conventional CT. Of these, three[115,120,121] found that there were no significant differences between the doses from the two versions of this technology. Two[114,116] measured a lower effective dose from SCT than from conventional CT, and one[119] determined that the skin dose was higher for SCT than for conventional CT.

There were also comparisons of SCT with five-view radiography[117] (which found a lower dose for SCT), and of dual and single slice SCT[118] (where the doses were the same).

The overall conclusion must be that the introduction of SCT will not increase the dose to this patient population unless the number of examinations increases.

EBCT

The only phantom study[124] of EBCT dose considered the dose to the breast for several examinations, including the investigation of CAC. The clinical studies that we included in this review for that examination also, in some cases, gave an estimate of dose, but this was usually expressed in terms of a total patient or skin dose. This was of the order of 10 mGy.

Changes in the knowledge base

Primary studies were included in this review only if they were published as full articles before January 1997. Although the full search and review protocol has not been performed on 1997 publications, we are aware of recent articles reporting findings of relevance to this review. In particular, more work on the use of SCT in PE has recently been published and includes several new studies.[238–240] Eleven ongoing studies were identified from the survey of experts at UK centres with access to SCT. Of these, two were applicable to the clinical applications covered in this review; one has since been published.[240] The other is a comparison of CT and MRI for imaging liver metastases.

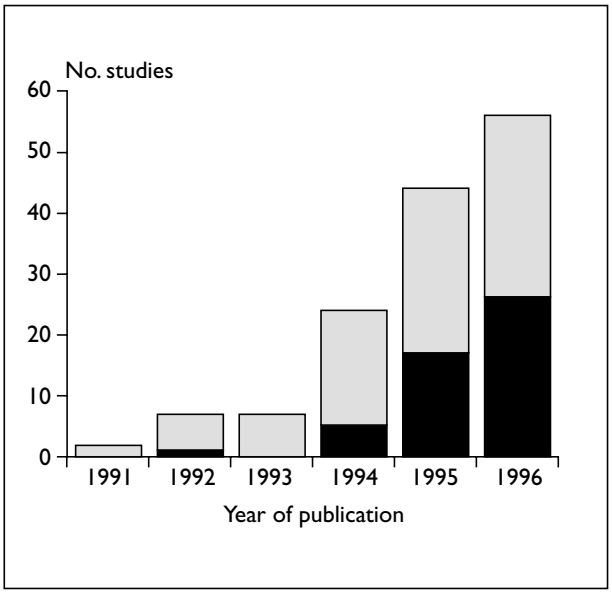


FIGURE 14 *Number of SCT studies included in and excluded from this review, plotted by year of publication (□, excluded; ■, included)*

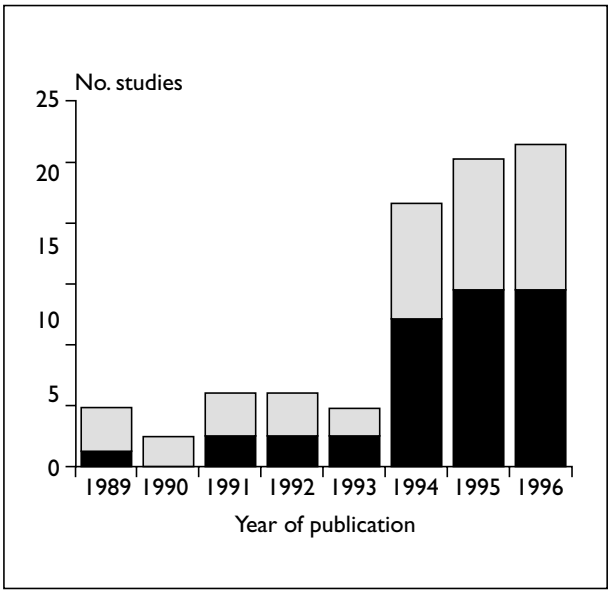


FIGURE 15 *Number of EBCT studies included in and excluded from this review, plotted by year of publication (□, excluded; ■, included)*

Figures 14 and *15* show the number of articles plotted by year for SCT and EBCT respectively. The dark bar is the number satisfying the inclusion criteria of this review; the lighter bar represents those excluded from the review. The bars together represent the total number of studies addressing questions in any of the hierarchical categories.

For SCT, there has been a steady increase in the number of published studies. There is no indication of reaching a plateau or of a sudden increase or decrease in these publications. From the trend, about 80 articles per annum will need to be retrieved and undergo the assessment process over the next 3 years. It can be seen that the proportion of studies satisfying the inclusion criteria set by this review remained about the same from 1994 to 1996.

For EBCT, the increase in the number of published studies is much less dramatic than for SCT and, indeed, the number satisfying the inclusion criteria may even be reaching a plateau. About 30 articles per annum will need to be retrieved and undergo the assessment process for the next 3 years. However, for the studies concerning the use of EBCT in asymptomatic groups to predict CAD, one would expect the number to rise in the future, once a reasonable length of follow-up time has elapsed.

In chapter 9 we comment on the difficulties of assessing evolving technologies. It is worth noting here that the CT technologies have not only evolved in the past; they will continue to evolve in the future. In the case of SCT, the acquisition time is expected to fall to 500 ms within the next 2 years, bringing it close to the 100 ms value for EBCT. This will introduce new clinical applications, such as the detection of coronary calcification.[241] In the case of EBCT, evolution will improve the spatial resolution beyond its current level of 10 line pairs per millimetre.

Chapter 9
Conclusion

Review methodology

Timeliness of the review

As far as SCT is concerned, our review justified the concern expressed in the commissioning brief, that the technology might come into general use without full evaluation. This has happened but, with hindsight, this is not surprising. The introduction of SCT was very much an example of a technology-driven advance. The devices could be viewed as simply the latest model CT machines, which one would automatically buy to replace a conventional machine without necessarily intending to make full use of the spiral capabilities. They had other improvements over the machines they would replace, in terms of improved user interface, rapid computing and the most recent interfacing capabilities. Health technology assessment performed to direct purchasing decisions would have to have been commissioned and completed within a 2-year period after the introduction of the technological advance, which in itself rules out any studies with long-term follow-up. Even with the benefit of hindsight, it is difficult to envisage an effective strategy for the timing of health technology assessment and delaying the spread of this evolving, commercially-driven technology.

Although the day has passed when it might have been possible to declare that there was no case for the purchase of an SCT device, it is not too late for advice on best practice to be provided; but it is unlikely that this will be in the form of a comparison with conventional scanning. The relevant applications are those that could not be performed on a conventional machine.

The situation for EBCT is different because these devices are not perceived as the latest model of an existing technology; nor were the benefits so immediately apparent that demand from anecdotal evidence caused widespread implementation. In the USA they were advocated for unproven applications soon after their introduction, but in the UK potential purchasers are in a position to benefit from the early health technology assessment reported in this review. The timing appears to have been good in this case.

Implications for searching

MEDLINE and BIDS-ISI provided all but one of the primary studies included in our EBCT review; in the SCT area, all of the articles were in one or both of these electronic databases. No further articles were found, in spite of the thorough, time-consuming search employed. As timeliness is so important in the systematic reviewing field, a methodological study to assess the impact of excluding studies that are not on the major databases would be of interest, to determine whether an approach following systematic principles, but relying on the major databases alone, gives valid results. Sufficient systematic reviews have now been carried out under the Health Technology Assessment Programme for a retrospective analysis across several subjects to be feasible. In both this and a previous medical imaging review[31] all the health economics papers were found in MEDLINE and BIDS-ISI. The economics papers are more easily identified than those at the patient outcome level, for example. Thus, relatively speedy health economics reviews may be feasible.

Study validity in medical imaging

As in our previous review,[31] the assessment of study validity was problematic. At the levels of patient outcome, therapeutic impact and diagnostic impact, the few studies that were found had a range of different designs and incomplete reporting of methodology. As a result, no attempt was made to apply a checklist approach to these studies. A very simple checklist was appropriate for health economics. At the diagnostic performance level, it was possible to apply a checklist, but difficulties were then experienced in trying to develop an objective validity scoring system based on the results. In one of our topic areas it was possible to perform a regression analysis that investigated the impact of study design features on the results it gave for diagnostic performance. At the significance level equivalent to $p < 0.05$ after application of the Bonferroni correction, no features were significant. This was partly related to widespread omissions in the reports of studies that made it hard to determine exactly how a study had been performed, and what its strengths and weaknesses were. It would be sensible for a consensus view to be agreed

about assessing study validity between those undertaking reviews in areas where randomised controlled trials do not exist. This would prevent numerous different methodologies being introduced and might result in a single robust one being agreed. There is even scope for comparative trials to be performed using results from existing reviews. This might be coordinated by the Cochrane Methods Working Group on Systematic Reviews of Screening and Diagnostic Tests.[41]

Further research is recommended into the suitability of SROC methodology for applications for which sensitivity and specificity differ widely. Of particular interest is the effect of sample size.

An awareness of the requirements for evidence-based diagnostic imaging is increasing, thanks to the efforts of professional organisations and the publication of review articles.[242,243] However, the major recommendation arising from this review is that, to ensure study validity and to facilitate the synthesis of results, imaging scientists must be encouraged both to perform better designed studies and to ensure that descriptions published in the literature are comprehensive.

Health economics

It was noted in chapter 8 that the only currently feasible approach to the economic evaluation of latest generation CT is decision modelling.

Decision modelling is useful only if the data sources are reliable. Considerable improvement is necessary in the quality of data collected in different types of study (e.g. cost, impact and outcome) before good evaluations of latest generation CT could be carried out using this approach. It would be unrealistic to expect the conduct of evaluations of diagnostic technologies to be transformed overnight to meet all the criteria specified in Drummond *et al.*[34] or the *BMJ*.[12] This should be regarded as a long-term objective. In the meantime, practical steps can be taken to improve the quality of information without a major additional burden on investigators.

For example, in estimating costs, the application of some simple principles could bring about a great improvement. The impact of the use of a diagnostic procedure on the use of healthcare resources should be measured comprehensively, including the impact on the use of other tests and the treatment ultimately given. Many diagnostic technologies involve significant capital investment. It is vital to take account of this when distinguishing between marginal and average costs over different periods. Too much reliance should not be placed on hospital charges as a source of unit costs. Charges are at best a proxy for average costs and are subject to systematic bias (addition of a profit margin) and random bias (differential charging in response to market competition). As the timescale of studies is extended, the importance of discounting costs must be understood.

Although the use of patient-based outcome measures in clinical studies is increasing, it is by no means universal, so success in this endeavour is not guaranteed. It is therefore important that as many new studies as possible in the diagnostic field should include outcome measures that can ultimately be used in economic evaluations.

Patient outcome

There is a lack of evidence at the patient outcome level and a need for new studies. Sufficiently extended follow-up to observe outcome must be included and the QoL indicators described in chapter 1 used. Any data at this level are also of value in full economic assessments, as discussed above.

Therapeutic impact

There is a lack of evidence at the therapeutic impact level and a need for more studies. Sufficiently extended follow-up must be included to observe changes in management and to avoid the problems described in chapter 1, which occur if only the therapy that actually was given is recorded. A design recording intentions to treat independently for the compared technologies is recommended. Any data at this level are also of value in full economic assessments, as discussed above. Based on the lack of evidence on therapeutic impact, an educational initiative would be justified to encourage medical imaging investigators intending to work at the diagnostic performance level to enhance their study protocol to include the assessment of therapeutic impact.

Diagnostic impact

SCT

There is no strong evidence available from comparative studies. Studies designed as suggested in chapter 1 are required, but it is no longer

relevant to compare SCT with conventional CT. Comparisons should take place with other modalities.

EBCT

The number of completed studies is low, but the indication is that the sensitivity of EBCT for diagnosing angiographic CAD is higher than for other techniques, but its specificity is low. Thus it cannot, in its present form, be recommended for this role. This is discussed further under diagnostic performance.

Diagnostic performance

There is a need to update this review in two of the clinical applications areas. First, the literature on the use of SCT for the diagnosis of PE has recently expanded and, secondly, that on the use of EBCT on asymptomatic individuals needs revisiting when studies with longer-term follow-up for subsequent cardiac events may be available.

SCT

HAP versus PVP

There is weak evidence that the HAP protocol provides additional information compared with the use of the PVP alone.

HAP versus reference

The evidence in this area is weak. A multicentre study, in which a single reference standard is used and the same acquisition protocol applied, would establish a baseline figure that is currently not available.

PVP protocols

It is not clear from the available literature what values for parameters should be recommended for the visualisation of hypovascular lesions because there was disagreement between the results of the included studies. A long delay time (of the order of 70 s) and a contrast medium dose of over 300 mg/ml in 125 ml seem to be advocated, but more studies are necessary. However, the reported success of automated protocols suggests that these will be standard in the future, so we recommend the following investigations.

- A multicentre study on a group of affected patients. This would look at the maximal parenchymal enhancement for a range of automatic protocols. The study would be tightly controlled with respect to intrinsic factors, including patient weight.[76]
- Studies on reducing the contrast medium dose. These would focus on lesion detectability rather than maximal parenchymal enhancement and would again use automatic protocols.

PE

As was stated in chapter 8, although there is uncertainty about the true sensitivity and specificity achievable, the future use of SCT in this clinical application depends upon the determination of the clinical importance of subsegmental PE. A systematic review of the literature on this topic is recommended. It is unlikely that clinical studies are the optimum means of determining the best diagnostic work-up for acute PE, and it was certainly not possible to draw any conclusions from the literature reviewed. Further research using decision analytical modelling should follow the review of subsegmental PE, comparing a variety of diagnostic strategies including VP, SCT, MRI and pulmonary angiography.

EBCT

EBCT prediction of asymptomatic CAD

At present there is no evidence that supports the use of EBCT in an asymptomatic population for the prediction of subsequent coronary events.

EBCT diagnosis of symptomatic CAD

The evidence in the literature suggests that the most appropriate use of EBCT is to employ the absence of CAC to exclude obstructive CAD in the older population. The available evidence is from rather mixed populations and the studies were not rigorous with their blinding protocols. There is scope for further work here.

If EBCT is used to detect atherosclerotic disease by the presence of CAC, the specificity of the test is low against the gold standard of angiography. This corresponds with a low PPV.

The choice of threshold for the diagnosis of a positive CAC finding appears to be inappropriate for use in a symptomatic population. As a result, the specificity of the test is low. Investigation into the use of higher thresholds is recommended. Where comparisons are to be made with angiography, the diagnosis of CAD on the angiogram should be made by considering both stenosis and lumen irregularity.

EBCT reproducibility for CAC

EBCT measurements of CAC are not sufficiently reproducible for use in longitudinal studies.

Radiation dose

SCT

The evidence is inconclusive, but indicates that the dose from SCT will not increase the dose to the population over that for conventional CT. Dose remains significant and future decision analytical modelling work should include radiation dose in the model, especially when comparison with non-ionising radiation modalities is made.

EBCT

There is very little evidence on which to base conclusions. A study to compare the dose with that from conventional coronary angiography is required if EBCT should become a preliminary to that investigation.

Dissemination and further research

The important target audiences for dissemination of the results of this review are the purchasing decision makers in the NHS and potential users of the technology (radiologists, cardiologists and surgeons). In the absence of sound economic data, no firm recommendations can be made to purchasers.

In terms of diagnostic performance, there is currently insufficient evidence in almost every area that we studied. There would seem to be no contraindications to the purchase of SCT, but the recommendations are negative for EBCT. Present evidence does not justify the use of EBCT devices for screening asymptomatic individuals. The literature should be monitored for further work in this area. There is already sufficient evidence showing that they should not be used for longitudinal studies of CAC.

From the literature reviewed, a number of study design faults were found to be particularly common. The new studies that we have recommended for diagnostic performance and impact should be designed to:

- use a single reference standard, ideally the recognised gold standard
- avoid verification bias by ensuring the single reference test is applied to all participants
- ensure that observers are blinded to the results of other studies, particularly to the reference result
- publish sufficient data for the completion of 2×2 contingency tables
- use published recommendations on sample size calculation to ensure that sufficient participants are included for statistical validity
- comment on operator dependence/ learning curves
- publish study design information to allow proper assessment of study quality.

In summary, the lessons learned from the health economics section of this review suggest the following strategy for economic evaluation in the diagnostic performance field:

- clarify diagnostic accuracy from good quality studies
- establish diagnostic and therapeutic impact where possible from such studies
- encourage new studies specifically to measure diagnostic impact, therapeutic impact and patient outcome
- encourage the use of patient-based outcome measures
- estimate costs from good quality studies
- use decision-modelling techniques to combine outcome and cost data from the different sources.

The final target audience is the entire medical imaging community, including those who perform studies, write and referee articles, and edit journals. A key point noted in this review was the poor quality of the written descriptions of published studies. It was not possible properly to assess study design because the pertinent information was missing from the descriptions. In the interests both of facilitating secondary research and of improving the quality of primary studies, the community must be made aware of the importance not only of designing a study well but of reporting the features of that design comprehensively.

Acknowledgements

The authors thank the following people for their invaluable contributions to the review.

External panel member:
Professor Ian Isherwood

Statistical support:
Sub-Unit for Medical Statistics (SUMS),
University of Leeds

Literature searching assistance:
Medical and Dental Library Staff,
University of Leeds
Ms K Saunders, British Institute of Radiology

Secretarial support:
Mrs J Pemberton, University of Leeds
Mr PE Lines

Respondents to information requests:
Professor JE Adams,
Manchester Royal Infirmary, UK

Professor DJ Allison,
Hammersmith Hospital, London, UK

Mr M Andrews,
Philips Medical Systems

Professor JJK Best,
Royal Infirmary of Edinburgh, UK

Dr RM Blaquiere,
Southampton General Hospital, UK

Professor G Cherryman,
Leicester Royal Infirmary, UK

Professor P Dawson,
Hammersmith Hospital, London, UK

Professor AK Dixon,
Addenbrooke's Hospital, Cambridge, UK

Dr D Hadley,
Southern General Hospital, Glasgow, UK

Mr M Haydon,
Philips Medical Systems

Professor JE Husband,
Royal Marsden Hospital, Surrey, UK

Professor A Jackson,
University of Manchester, UK

Dr PC Jackson,
Southampton General Hospital, UK

Dr EM McIlrath,
Royal Victoria Hospital, Belfast, UK

Dr JP Owen,
Royal Victoria Infirmary, Newcastle upon Tyne, UK

Professor GM Roberts,
University of Wales College of Medicine, Cardiff, UK

Mr J Stoddart,
Siemens Medical Engineering

Professor J Weir,
Aberdeen Royal Infirmary, UK

Ms W Wetherfield,
Toshiba Medical Systems

Professor GH Whitehouse,
University of Liverpool, UK

Professor BS Worthington,
Queen's Medical Centre, Nottingham, UK

Translation:
Ms AN Rodriguez

This work was commissioned by the NHS Health Technology Assessment Programme, project 94/28/01. The views and opinions expressed do not necessarily reflect those of The Secretary of State for Health. In part, this work was undertaken by the Leeds Teaching Hospitals NHS Trust, who received funding from the NHS Executive; the views expressed in this publication are those of the authors and not necessarily those of the NHS Executive.

References

1. Mowatt G, Bower DJ, Brebner JA, Cairns JA, Grant AM, McKee L. When and how to assess fast-changing technologies: a comparative study of medical applications of four generic technologies. *Health Technol Assess* 1997;**1**(14).
2. Kalender WA. Principles and applications of spiral CT. *Nucl Med Biol* 1994;**21**:693–9.
3. McCollough CH, Morin RL. The technical design and performance of ultrafast computed tomography. *Radiol Clin North Am* 1994;**32**:521–36.
4. Kalender WA, Polacin A, Suss C. A comparison of conventional and spiral CT: an experimental study on the detection of spherical lesions. *J Comput Assist Tomogr* 1994;**18**:167–76.
5. Stehling MK, Lawrence JA, Weintraub JL, Raptopoulos V. CT angiography: expanded clinical applications. *AJR Am J Roentgenol* 1994;**163**:947–55.
6. Fineberg HV, Bauman R, Sosman M. Computerized cranial tomography. Effect on diagnostic and therapeutic plans. *JAMA* 1977;**238**:224–7.
7. Mackenzie R, Dixon AK. Measuring the effects of imaging: an evaluative framework. *Clin Radiol* 1995;**50**:513–18.
8. Thornbury JR. Clinical efficacy of diagnostic imaging: love it or leave it. *AJR Am J Roentgenol* 1994;**162**:1–8.
9. Spilker B, editor. Quality of life and pharmacoeconomics in clinical trials. 2nd ed. Philadelphia, PA: Lippincott-Raven; 1996.
10. Guyatt GH, Naylor CD, Juniper E, Heyland DK, Jaeschke R, Cook DJ. Users' guides to the medical literature: XII. How to use articles about health-related quality of life. Evidence-Based Medicine Working Group. *JAMA* 1997;**277**:1232–7.
11. Fries J, Singh G. The hierachy of patient outcomes. In: Spilker B, editor. Quality of life and pharmacoeconomics in clinical trials. 2nd ed. Philadelphia: Lippincott-Raven; 1996:33–40.
12. Drummond MF, Jefferson TO. Guidelines for authors and peer reviewers of economic submissions to the *BMJ*. The BMJ Economic Evaluation Working Party. *BMJ* 1996;**313**:275–83.
13. Williams A. Economics of coronary artery bypass grafting. *BMJ (Clin Res)* 1985;**291**:326–9.
14. Bryan S. Economic evaluation of diagnostic technologies: a rule for conjoint analysis? *Soc Sci Health* 1997;**3**:209–11.
15. Kelly S, Berry E, Roderick P, Harris KM, Cullingworth J, Gathercole L, *et al.* The identification of bias in studies of the diagnostic performance of imaging modalities. *Br J Radiol* 1997;**70**:1028–35.
16. Oliver JH, Baron RL. Helical biphasic contrast-enhanced CT of the liver: technique, indications, interpretation, and pitfalls. *Radiology* 1996;**201**:1–14.
17. Robinson PJA. Ventilation–perfusion lung scanning and spiral computed tomography of the lungs: competing or complementary modalities? *Eur J Nucl Med* 1996;**23**:1547–53.
18. The PIOPED Investigators. Value of the ventilation/perfusion scan in acute pulmonary embolism. Results of the prospective investigation of pulmonary embolism diagnosis (PIOPED). *JAMA* 1990;**263**:2753–9.
19. van Erkel A, van Rossum A, Bloem JL, Kievit J, Pattynama PM. Spiral CT angiography for suspected pulmonary embolism: a cost-effectiveness analysis. *Radiology* 1996;**201**:29–36.
20. Remy-Jardin M, Remy J, Wattinne L, Giraud F. Central pulmonary thromboembolism: diagnosis with spiral volumetric CT with the single-breath-hold technique – comparison with pulmonary angiography. *Radiology* 1992;**185**:381–7.
21. Bleiweis MS, Georgiou D, Brundage BH. Ultrafast CT and the cardiovascular system. *Int J Card Imaging* 1992;**8**:289–302.
22. Agatston AS, Janowitz WR, Hildner FJ, Zusmer NR, Viamonte MJ, Detrano R. Quantification of coronary artery calcium using ultrafast computed tomography. *J Am Coll Cardiol* 1990;**15**:827–32.
23. Office of Population Censuses and Surveys. Mortality statistics 1995. (DH2.) London: HMSO, 1997:22.
24. Department of Health. Health of the nation briefing pack. 2nd ed. A strategy for health in England. London: DoH, 1997. Also available at ftp://ftp.open.gov.uk/pub/docs/doh/hotnpack.pdf
25. Wong ND, Detrano RC, Abrahamson D, Tobis JM, Gardin JM. Coronary artery screening by electron beam computed tomography. Facts, controversy, and future. *Circulation* 1995;**92**:632–6.
26. Viamonte MJ, Janowitz W, Agatston A. What ultrafast CT (UFCT) can and cannot do in coronary atherosclerosis. *Cathet Cardiovasc Diagn* 1993;**30**:180–1.

27. Wexler L, Brundage B, Crouse J, Detrano R, Fuster V, Maddahi J, *et al.* Coronary artery calcification: pathophysiology, epidemiology, imaging methods, and clinical implications. A statement for health professionals from the American Heart Association Writing Group. *Circulation* 1996;**94**:1175–92.

28. NHS Centre for Reviews and Dissemination. Undertaking systematic reviews of research on effectiveness (CRD Report 4). York: University of York, 1996.

29. The British Library. Inside Information Plus. 1998: http://www.bl.uk/online/inside

30. Online Computer Library Center. 1998: http://www.oclc.org

31. Harris KM, Kelly S, Berry E, Hutton J, Roderick P, Cullingworth J, *et al.* Systematic review of endoscopic ultrasound in gastro-oesophageal cancer. *Health Technol Assess* 1998;**2**(18).

32. Williams A. The cost–benefit approach. *Br Med Bull* 1974;**30**:252–6.

33. McMaster University. How to read clinical journals: VII. To understand an economic evaluation (part B). *Can Med Assoc J* 1984;**130**:1542–9.

34. Drummond MF, Stoddart GL, Torrance GW. Methods for the economic evaluation of health care programmes. Oxford: Oxford University Press, 1997.

35. Udvarhelyi IS, Colditz GA, Rai A, Epstein AM. Cost-effectiveness and cost–benefit analyses in the medical literature. Are the methods being used correctly? *Ann Intern Med* 1992;**116**:238–44.

36. Adams ME, McCall NT, Gray DT, Orza MJ, Chalmers TC. Economic analysis in randomized controlled trials. *Med Care* 1992;**30**:231–43.

37. Hutton J, Clark M, Sanderson D. The cost-effectiveness of MRI in the DGH: a review. *Health Technol Assess* (In preparation.)

38. Mulrow CD, Linn WD, Gaul MK, Pugh JA. Assessing quality of a diagnostic test evaluation. *J Gen Intern Med* 1989;**4**:288–95.

39. Jaeschke R, Guyatt G, Sackett DL. Users' guides to the medical literature: III. How to use an article about a diagnostic test. A. Are the results of the study valid? Evidence-Based Medicine Working Group. *JAMA* 1994;**271**:389–91.

40. Jaeschke R, Guyatt GH, Sackett DL. Users' guides to the medical literature: III. How to use an article about a diagnostic test. B. What are the results and will they help me in caring for my patients? Evidence-Based Medicine Working Group. *JAMA* 1994;**271**:703–7.

41. Cochrane Methods Working Group on Systematic Reviews of Screening and Diagnostic Tests. Recommended method, updated 6 June 1996. http://som.flinders.edu.au/FUSA/COCHRANE/cochrane/sadtdoc1.htm

42. Berry E, Kelly S, Harris KM, Cullingworth J, Gathercole L, Hutton J, *et al.* Development of a methodology for systematic literature reviews in diagnostic imaging [Abstract]. *Br J Radiol* 1997;**70**(Suppl):109.

43. Oxman AD, Guyatt GH. Validation of an index of the quality of review articles. *J Clin Epidemiol* 1991;**44**:1271–8.

44. Metz CE. Some practical issues of experimental design and data analysis in radiological ROC studies. *Invest Radiol* 1989;**24**:234–45.

45. Moses LE, Shapiro D, Littenberg B. Combining independent studies of a diagnostic test into a summary ROC curve: data-analytic approaches and some additional considerations. *Stat Med* 1993;**12**:1293–316.

46. Irwig L, Tosteson AN, Gatsonis C, Lau J, Colditz G, Chalmers TC, *et al.* Guidelines for meta-analyses evaluating diagnostic tests. *Ann Intern Med* 1994;**120**:667–76.

47. Irwig L, Macaskill P, Glasziou P, Fahey M. Meta-analytic methods for diagnostic test accuracy. *J Clin Epidemiol* 1995;**48**:119–30.

48. Kaneko M, Eguchi K, Ohmatsu H, Kakinuma R, Naruke T, Suemasu K, *et al.* Peripheral lung cancer: screening and detection with low-dose spiral CT vs radiography. *Radiology* 1996;**201**:798–802.

49. Lindgren B, Demos T, Marson R, Posniak H, Kostro B, Calvert D, *et al.* Renal computed tomography with 3-dimensional angiography and simultaneous measurement of plasma contrast clearance reduce the invasiveness and cost of evaluating living renal donor candidates. *Transplantation* 1996;**61**:219–23.

50. Semelka R, Schlund J, Molina P, Willms AB, Kahlenberg M, Mauro MA, *et al.* Malignant liver lesions: comparison of spiral CT arterial portography and MR imaging for diagnostic accuracy, cost, and effect on patient management. *J Magn Reson Imaging* 1996;**1**:39–43.

51. Hernigou A, Perrin JP, Grataloup C, Philippe E, Plainfosse MC. Cost comparison of electron beam tomography with conventional computed tomography scanning. *Acad Radiol* 1996;**3**:5145–6.

52. Stanford W, Travis ME, Thompson BH, Reiners TJ, Hasson RR, Winniford MD. Electron beam computed tomographic detection of coronary calcification in patients undergoing percutaneous transluminal coronary angioplasty: predictability of restenosis. A preliminary report. *Am J Card Imaging* 1995;**9**:257–60.

53. Garden AS, Morrison WH, Clayman GL, Ang KK, Peters LJ. Early squamous cell carcinoma of the hypopharynx: outcomes of treatment with radiation alone to the primary disease. *Head Neck* 1996;**18**:317–22.

54. Wong ND, Detrano RC, Diamond G, Rezayat C, Mahmoudi R, Chong EC, *et al.* Does coronary artery screening by electron beam computed tomography motivate potentially beneficial lifestyle behaviors? *Am J Cardiol* 1996;**78**:1220–3.

55. Lacrosse M, Trigaux JP, van Beers BE, Weynants P. 3D spiral CT of the tracheobronchial tree. *J Comput Assist Tomogr* 1995;**19**:341–7.

56. Winston CB, Wechsler RJ, Salazar AM, Kurtz AB, Spirn PW. Incidental pulmonary emboli detected at helical CT: effect on patient care. *Radiology* 1996;**201**:23–7.

57. Barloon TJ, Galvin JR, Mori M, Stanford W, Gingrich RD. High-resolution ultrafast chest CT in the clinical management of febrile bone marrow transplant patients with normal or nonspecific chest roentgenograms. *Chest* 1991;**99**:928–33.

58. Mousseaux E, Hernigou A, Azencot M, Sapoval M, Auguste M, Gaux JC. Evaluation by electron beam computed tomography of intracardiac masses suspected by transoesophageal echocardiography. *Heart* 1996;**76**:256–63.

59. Szolar DH, Groell R, Braun H, Preidler K, Stiskal M, Kern R, *et al.* Ultrafast computed tomography and three-dimensional image processing of CT sialography in patients with parotid masses poorly defined by magnetic resonance imaging. *Acta Otolaryngol* 1996;**116**:112–18.

60. Jung G, Krahe T, Krug B, Hahn U, Raab M. Delineation of segmental liver anatomy. Comparison of ultrasonography, spiral CT and MR imaging for preoperative localization of focal liver lesions to specific hepatic segments. *Acta Radiol* 1996;**37**:691–5.

61. Oi H, Murakami T, Kim T, Matsushita M, Kishimoto H, Nakamura H. Dynamic MR imaging and early-phase helical CT for detecting small intrahepatic metastases of hepatocellular carcinoma. *AJR Am J Roentgenol* 1996;**166**:369–74.

62. Ueda K, Kitagawa K, Kadoya M, Matsui O, Takashima T, Yamahana T. Detection of hypervascular hepatocellular carcinoma by using spiral volumetric CT: comparison of US and MR imaging. *Abdom Imaging* 1995;**20**:547–53.

63. van Hoe L, Bosmans H, Aerts P, Baert AL, Fevery J, Kiefer B, *et al.* Focal liver lesions: fast T2-weighted MR imaging with half-Fourier rapid acquisition with relaxation enhancement. *Radiology* 1996;**201**:817–23.

64. Yamashita Y, Mitsuzaki K, Yi T, Ogata I, Nishiharu T, Urata J, *et al.* Small hepatocellular carcinoma in patients with chronic liver damage: prospective comparison of detection with dynamic MR imaging and helical CT of the whole liver. *Radiology* 1996;**200**:79–84.

65. Bonaldi VM, Bret PM, Reinhold C, Atri M. Comparison of helical and conventional computed tomography of the liver. *Can Assoc Radiol J* 1995;**46**:443–8.

66. Choi BI, Shin YM, Han JK, Chung JW, Park JH, Han MC. Focal hepatic nodules after transcatheter oily chemoembolization: detection with spiral CT versus conventional CT. *Abdom Imaging* 1996;**21**:33–6.

67. Fujita M, Kuroda C, Kumatani T, Yoshioka H, Kuriyama K, Inoue E, *et al.* Comparison between conventional and spiral CT in patients with hypervascular hepatocellular carcinoma. *Eur J Radiol* 1994;**18**:134–6.

68. Polger M, Seltzer SE, Head BL, Savci G, Silverman SG, Adams DF. Spiral computed tomography of the liver – contrast agent pharmacokinetics and the potential for improved hepatic enhancement. *Acad Radiol* 1995;**2**:19–25.

69. Remy-Jardin M, Remy J, Deschildre F, Artaud D, Beregi JP, Hossein-Foucher C, *et al.* Diagnosis of pulmonary embolism with spiral CT: comparison with pulmonary angiography and scintigraphy. *Radiology* 1996;**200**:699–706.

70. Kajinami K, Seki H, Takekoshi N, Mabuchi H. Noninvasive prediction of coronary atherosclerosis by quantification of coronary artery calcification using electron beam computed tomography – comparison with electrocardiographic and thallium exercise stress test results. *J Am Coll Cardiol* 1995;**26**:1209–21.

71. Oliver JH, Baron RL, Federle MP, Rockette HEJ. Detecting hepatocellular carcinoma: value of unenhanced or arterial phase CT imaging or both used in conjunction with conventional portal venous phase contrast-enhanced CT imaging. *AJR Am J Roentgenol* 1996;**167**:71–7.

72. Bonaldi VM, Bret PM, Reinhold C, Atri M. Helical CT of the liver: value of an early hepatic arterial phase. *Radiology* 1995;**197**:357–63.

73. Hollett MD, Jeffrey RBJ, Nino-Murcia M, Jorgensen MJ, Harris DP. Dual-phase helical CT of the liver: value of arterial phase scans in the detection of small (< or = 1.5 cm) malignant hepatic neoplasms. *AJR Am J Roentgenol* 1995;**164**:879–84.

74. Mitsuzaki K, Yamashita Y, Ogata I, Nishiharu T, Urata J, Takahashi M. Multiple-phase helical CT of the liver for detecting small hepatomas in patients with liver cirrhosis: contrast-injection protocol and optimal timing. *AJR Am J Roentgenol* 1996;**167**:753–7.

75. Choi BI, Cho JM, Han JK, Choi DS, Han MC. Spiral CT for the detection of hepatocellular carcinomas: relative value of arterial- and late-phase scanning. *Abdom Imaging* 1996;**21**:440–4.

76. Dodd GD, Baron RL. Investigation of contrast enhancement in CT of the liver: the need for improved methods. *AJR Am J Roentgenol* 1993;**160**:643–6.

77. Brink JA, Heiken JP, Forman HP, Sagel SS, Molina PL, Brown PC. Hepatic spiral CT: reduction of dose of intravenous contrast material. *Radiology* 1995;**197**:83–8.

78. Herts BR, Paushter DM, Einstein DM, Zepp R, Friedman RA, Obuchowski N. Use of contrast material for spiral CT of the abdomen: comparison of hepatic enhancement and vascular attenuation for three different contrast media at two different delay times. *AJR Am J Roentgenol* 1995;**164**:327–31.

79. Hoeffel C, Legmann P, Girardot C, Bunouf P, Mezzi K, Bonnin A. Contrast material injection for hepatic helical computed tomography. Comparative study of five protocols. *Invest Radiol* 1996;**31**:467–71.

80. Kopka L, Funke M, Vosshenrich R, Hagemann A, Oestmann JW, Grabbe E. Helical CT of the liver: evaluation of injection flow rate, mode, and scan delay with a reduced-volume contrast medium bolus. *J Comput Assist Tomogr* 1995;**19**:406–11.

81. Silverman PM, O'Malley J, Tefft MC, Cooper C, Zeman RK. Conspicuity of hepatic metastases on helical CT: effect of different time delays between contrast administration and scanning. *AJR Am J Roentgenol* 1995;**164**:619–23.

82. Small WC, Nelson RC, Bernardino ME, Brummer LT. Contrast-enhanced spiral CT of the liver: effect of different amounts and injection rates of contrast material on early contrast enhancement. *AJR Am J Roentgenol* 1994;**163**:87–92.

83. Kopka L, Funke M, Fischer U, Vosshenrich R, Oestmann JW, Grabbe E. Parenchymal liver enhancement with bolus-triggered helical CT: preliminary clinical results. *Radiology* 1995;**195**:282–4.

84. Kopka L, Rodenwaldt J, Fischer U, Mueller DW, Oestmann JW, Grabbe E. Dual-phase helical CT of the liver: effects of bolus tracking and different volumes of contrast material. *Radiology* 1996;**201**:321–6.

85. Silverman PM, Roberts S, Tefft MC, Brown B, Fox SH, Cooper C, *et al.* Helical CT of the liver: clinical application of an automated computer technique, SmartPrep, for obtaining images with optimal contrast enhancement. *AJR Am J Roentgenol* 1995;**165**:73–8.

86. Silverman PM, Roberts SC, Ducic I, Tefft MC, Olson MC, Cooper C, *et al.* Assessment of a technology that permits individualized scan delays on helical hepatic CT: a technique to improve efficiency in use of contrast material. *AJR Am J Roentgenol* 1996;**167**:79–84.

87. Goodman LR, Curtin JJ, Mewissen MW, Foley WD, Lipchik RJ, Crain MR, *et al.* Detection of pulmonary embolism in patients with unresolved clinical and scintigraphic diagnosis: helical CT versus angiography. *AJR Am J Roentgenol* 1995;**164**:1369–74.

88. van Rossum AB, Pattynama PM, Ton ER, Treurniet FE, Arndt JW, van Eck B, *et al.* Pulmonary embolism: validation of spiral CT angiography in 149 patients. *Radiology* 1996;**201**:467–70.

89. Arad Y, Spadaro LA, Goodman K, Lledoperez A, Sherman S, Lerner C, *et al.* Predictive value of electron beam computed tomography of the coronary arteries – 19-month follow-up of 1173 asymptomatic subjects. *Circulation* 1996;**93**:1951–3.

90. Coin CG. Noninvasive visualization of coronary arteriosclerosis by computed tomography. *South Med J* 1993;**86**:295–6.

91. Barbir M, Bowker T, Ludman PF, Mitchell AG, Wood D, Yacoub M. Ultrafast computed tomographic scanning for detection of coronary disease in cardiac transplant recipients. *Am J Cardiol* 1994;**74**:941–4.

92. Bielak LF, Kaufmann RB, Moll PP, McCollough CH, Schwartz RS, Sheedy PF. Small lesions in the heart identified at electron beam CT: calcification or noise? *Radiology* 1994;**192**:631–6.

93. Bormann JL, Stanford W, Stenberg RG, Winniford D, Berbaum KS, Talman CL, *et al.* Ultrafast computed tomography detection of coronary artery calcification as an indicator of stenosis. *Am J Card Imaging* 1992;**6**:191–6.

94. Braun J, Oldendorf M, Moshage W, Heidler R, Zeitler E, Luft FC. Electron-beam computed tomography in the evaluation of cardiac calcifications in chronic dialysis patients. *Am J Kidney Dis* 1996;**27**:394–401.

95. Breen JF, Sheedy PF, Schwartz RS, Stanson AW, Kaufmann RB, Moll PP, *et al.* Coronary artery calcification detected with ultrafast CT as an indication of coronary artery disease. *Radiology* 1992;**185**:435–9.

96. Budoff MJ, Georgiou D, Brody A, Agatston AS, Kennedy J, Wolfkiel C, *et al.* Ultrafast computed tomography as a diagnostic modality in the detection of coronary artery disease: a multicenter study. *Circulation* 1996;**93**:898–904.

97. Devries S, Wolfkiel C, Fusman B, Bakdash H, Ahmed A, Levy P, *et al.* Influence of age and gender on the presence of coronary calcium detected by ultrafast computed tomography. *J Am Coll Cardiol* 1995;**25**:76–82.

98. Fallavollita JA, Brody AS, Bunnell IL, Kumar K, Canty JM. Fast computed tomography detection of coronary calcification in the diagnosis of coronary artery disease – comparison with angiography in patients less than 50 years old. *Circulation* 1994;**89**:285–90.

99. Rumberger JA, Sheedy PF, Breen JF, Schwartz RS. Coronary calcium, as determined by electron beam computed tomography, and coronary disease on arteriogram – effect of patient's sex on diagnosis. *Circulation* 1995;**91**:1363–7.

100. Tanenbaum SR, Kondos GT, Veselik KE, Prendergast MR, Brundage BH, Chomka EV. Detection of calcific deposits in coronary arteries by ultrafast computed tomography and correlation with angiography. *Am J Cardiol* 1989;**63**:870–2.

101. Yaghoubi S, Tang W, Wang S, Reed J, Hsiai J, Detrano R, *et al.* Offline assessment of atherosclerotic coronary calcium from electron beam tomograms. *Am J Card Imaging* 1995;**9**:231–6.

102. Agatston AS, Janowitz WR, Kaplan G, Gasso J, Hildner F, Viamonte M. Ultrafast computed tomography-detected coronary calcium reflects the angiographic extent of coronary arterial atherosclerosis. *Am J Cardiol* 1994;**74**:1272–4.

103. Kaufmann RB, Sheedy PF, Maher JE, Bielak LF, Breen JF, Schwartz RS, *et al.* Quantity of coronary artery calcium detected by electron beam computed tomography in asymptomatic subjects and angiographically studied patients. *Mayo Clin Proc* 1995;**70**:223–32.

104. Shields JP, Mielke CHJ, Watson P, Viren F. Comparison of 10, 20, and 40 level electron beam computed tomography studies for coronary calcium. *Am J Card Imaging* 1996;**10**:235–8.

105. Wong ND, Vo A, Abrahamson D, Tobis JM, Eisenberg H, Detrano RC. Detection of coronary artery calcium by ultrafast computed tomography and its relation to clinical evidence of coronary artery disease. *Am J Cardiol* 1994;**73**:223–7.

106. Devries S, Wolfkiel C, Shah V, Chomka E, Rich S. Reproducibility of the measurement of coronary calcium with ultrafast computed tomography. *Am J Cardiol* 1995;**75**:973–5.

107. Hernigou A, Challande P, Boudeville JC, Sene V, Grataloup C, Plainfosse MC. Reproducibility of coronary calcification detection with electron beam computed tomography. *Eur Radiol* 1996;**6**:210–16.

108. Kajinami K, Seki H, Takekoshi N, Mabuchi H. Quantification of coronary artery calcification using ultrafast computed tomography – reproducibility of measurements. *Coron Artery Dis* 1993;**4**:1103–8.

109. Kaufmann RB, Sheedy PF, Breen JF, Kelzenberg JR, Kruger BL, Schwartz RS, *et al.* Detection of heart calcification with electron beam CT – interobserver and intraobserver reliability for scoring quantification. *Radiology* 1994;**190**:347–52.

110. Shields JP, Mielke CHJ, Rockwood TH, Short RA, Viren FK. Reliability of electron beam computed tomography to detect coronary artery calcification. *Am J Card Imaging* 1995;**9**:62–6.

111. Shields JP, Mielke CHJ, Watson P. Inter-rater reliability of electron beam computed tomography to detect coronary artery calcification. *Am J Card Imaging* 1996;**10**:91–6.

112. Wang SJ, Detrano RC, Secci A, Tang WY, Doherty TM, Puentes G, *et al.* Detection of coronary calcification with electron beam computed tomography – evaluation of interexamination reproducibility and comparison of 3 image-acquisition protocols. *Am Heart J* 1996;**132**:550–8.

113. Bland J, Altman D. Statistical methods for assessing agreement between two methods of clinical measurement. *Lancet* 1986;**i**:307–10.

114. Chang LL, Chen FD, Chang PS, Liu CC, Lien HL. Assessment of dose and risk to the body following conventional and spiral computed tomography. *Chung Hua I Hsueh Tsa Chih* 1995;**55**:283–9.

115. Collie DA, Wright AR, Williams JR, Hashemi-Malayeri B, Stevenson AJ, Turnbull CM. Comparison of spiral-acquisition computed tomography and conventional computed tomography in the assessment of pulmonary metastatic disease. *Br J Radiol* 1994;**67**:436–44.

116. Jurik AG, Jensen LC, Hansen J. Radiation dose by spiral CT and conventional tomography of the sternoclavicular joints and the manubrium sterni. *Skeletal Radiol* 1996;**25**:467–70.

117. Jurik AG, Jensen LC, Hansen J. Total effective radiation dose from spiral CT and conventional radiography of the pelvis with regard to fracture classification. *Acta Radiol* 1996;**37**:651–4.

118. Liang Y, Kruger RA. Dual-slice spiral versus single-slice spiral scanning: comparison of the physical performance of two computed tomography scanners. *Med Phys* 1996;**23**:205–20.

119. Lucidarme O, Grenier P, Coche E, Lenoir S, Aubert B, Beigelman C. Bronchiectasis: comparative assessment with thin-section CT and helical CT. *Radiology* 1996;**200**:673–9.

120. McGhee PL, Humphreys S. Radiation dose associated with spiral computed tomography. *Can Assoc Radiol J* 1994;**45**:124–9.

121. Verdun FR, Meuli RA, Bochud FO, Imsand C, Raimondi S, Schnyder P, *et al.* Image quality and dose in spiral computed tomography. *Eur Radiol* 1996;**6**:485–8.

122. Levrero F, Coscia G, Pilot A, Cavagnetto F, Zucchi E. Image quality and dose distribution in multiscan and fast ring CT systems. *Radiat Prot Dosim* 1995;**57**:135–8.

123. McCollough CH, Liu HH. Breast dose during electron beam CT – measurement with film dosimetry. *Radiology* 1995;**196**:153–7.

124. McCollough CH, Zink FE, Morin RL. Radiation dosimetry for electron beam CT. *Radiology* 1994;**192**:637–43.

125. Balm R, Kaatee R, Blankensteijn JD, Mali WPTM, Eikelboom BC. CT-angiography of abdominal aortic aneurysms after transfemoral endovascular aneurysm management. *Eur J Vasc Endovasc Surg* 1996;**12**:182–8.

126. Brown CE, Purdy P, Cunningham FG. Head computed tomographic scans in women with eclampsia. *Am J Obstet Gynecol* 1988;**159**:915–20.

127. Costello P, Dupuy DE, Ecker CP, Tello R. Spiral CT of the thorax with reduced volume of contrast material: a comparative study. *Radiology* 1992;**183**:663–6.

128. Costello P, Gaa J. Spiral CT angiography of abdominal aortic aneurysms. *Radiographics* 1995;**15**:397–406.

129. Cunningham JD, Glajchen N, Brower ST. The use of spiral computed tomography in the evaluation of vessel encasement for pancreatic cancer. *Int J Pancreatol* 1996;**19**:9–14.

130. Engeler CE, Wasserman NF, Zhang G. Preoperative assessment of prostatic carcinoma by computerized tomography. Weaknesses and new perspectives. *Urology* 1992;**40**:346–50.

131. Gavant ML, Menke PG, Fabian T, Flick PA, Graney MJ, Gold RE. Blunt traumatic aortic rupture: detection with helical CT of the chest. *Radiology* 1995;**197**:125–33.

132. Helmberger H, Baum U, Dittler HJ, Sendler A, Schulte B, Herter B, *et al.* Adenocarcinoma of the gastroesophageal junction – CT for monitoring during neoadjuvant chemotherapy. *Eur J Radiol* 1996;**23**:107–10.

133. Hsieh J. A general approach to the reconstruction of X-ray helical computed tomography. *Med Phys* 1996;**23**:221–9.

134. Johnston DG, Hall K, Kendall-Taylor P, Ross WM, Crombie AL, Cook DB, *et al.* The long-term effects of megavoltage radiotherapy as sole or combined therapy for large prolactinomas: studies with high definition computerized tomography. *Clin Endocrinol* 1986;**24**:675–85.

135. Kato N, Semba CP, Dake MD. Embolization of perigraft leaks after endovascular stent–graft treatment of aortic aneurysms. *J Vasc Interv Radiology* 1996;**7**:805–11.

136. Kihara Y, Tamura S, Yuki Y, Kakitsubata S, Sugimura H, Kakitsubata Y, *et al.* Optimal timing for delineation of hepatocellular carcinoma in dynamic CT. *J Comput Assist Tomogr* 1993;**17**:719–22.

137. Robinson EK, Lee JE, Lowy AM, Fenoglio CJ, Pisters PWT, Evans DB. Reoperative pancreaticoduodenectomy for periampullary carcinoma. *Am J Surg* 1996;**172**:432–8.

138. Rubin GD. Three-dimensional helical CT angiography. *Radiographics* 1994;**14**:905–12.

139. Stevenson JC, Banks LM. Vertebral CT scanning. *Scott Med J* 1986;**31**:207–10.

140. Watanabe H, Nakata H, Egashira K, Nakamura K. Spiral volumetric CT as a routine technique for thoracic imaging. *J Thorac Imaging* 1993;**8**:316–20.

141. Wehmoller M, Eufinger H, Kruse D, Massberg W. CAD by processing of computed tomography data and CAM of individually designed prostheses. *Int J Oral Maxillofac Surg* 1995;**24**:90–7.

142. Winter TC, Freeny PC, Nghiem HV, Hommeyer SC, Barr D, Croghan AM, *et al.* Hepatic arterial anatomy in transplantation candidates: evaluation with three-dimensional CT arteriography. *Radiology* 1995;**195**:363–70.

143. Woodhouse CE, Ney DR, Sitzmann JV, Fishman EK. Spiral computed tomography arterial portography with 3-dimensional volumetric rendering for oncologic surgery planning – a retrospective analysis. *Invest Radiol* 1994;**29**:1031–7.

144. Berland LL. Screening for diffuse and focal liver disease: the case for hepatic computed tomography. *J Clin Ultrasound* 1984;**12**:83–9.

145. Conway WF, Hayes CW, Loughran T, Totty WG, Griffeth LK, Elkhoury GY, *et al.* Cross-sectional imaging of the patellofemoral joint and surrounding structures. *Radiographics* 1991;**11**:195–217.

146. Hamberg LM, Hunter GJ, Kierstead D, Lo EH, Gilberto GR, Wolf GL. Measurement of cerebral blood volume with subtraction three-dimensional functional CT. *Am J Neuroradiol* 1996;**17**:1861–9.

147. Koehler PR, Moss AA. Diagnosis of intra-abdominal and pelvic abscesses by computerized tomography. *JAMA* 1980;**244**:49–52.

148. Stern EJ, Webb WR, Gamsu G. Dynamic quantitative computed tomography. A predictor of pulmonary function in obstructive lung diseases. *Invest Radiol* 1994;**29**:564–9.

149. Wolf GL. Magnetic resonance imaging and the future of cardiac imaging. *Am J Cardiol* 1989; **64**:60E–3E.

150. Adachi H, Ino T, Mizuhara A, Yamaguchi A, Kobayashi Y, Nagai J. Assessment of aortic disease using three-dimensional CT angiography. *J Card Surg* 1994;**9**:673–8.

151. Gerber JD, Ney DR, Magid D, Fishman EK. Simulated femoral repositioning with three-dimensional CT. *J Comput Assist Tomogr* 1991;**15**:121–5.

152. Kauczor HU, Schwickert HC, Mayer E, Schweden F, Schild HH, Thelen M. Spiral CT of bronchial arteries in chronic thromboembolism. *J Comput Assist Tomogr* 1994;**18**:855–61.

153. LoCicero J, Costello P, Campos CT, Francalanci N, Dushay KM, Silvestri RC, *et al.* Spiral CT with multiplanar and three-dimensional reconstructions accurately predicts tracheobronchial pathology. *Ann Thorac Surg* 1996;**62**:818–22; discussion 822–3.

154. McEnery KW, Wilson AJ, Pilgram TK, Murphy WAJ, Marushack MM. Fractures of the tibial plateau: value of spiral CT coronal plane reconstructions for detecting displacement in vitro. *AJR Am J Roentgenol* 1994;**163**:1177–81.

155. Nunez DJ, Becerra JL, Fuentes D, Pagson S. Traumatic occlusion of the renal artery: helical CT diagnosis. *AJR Am J Roentgenol* 1996;**167**:777–80.

156. Padhani AR, Fishman EK, Heitmiller RF, Wang KP, Wheeler JH, Kuhlman JE. Multiplanar display of spiral CT data of the pulmonary hila in patients with lung cancer. Preliminary observations. *Clin Imaging* 1995;**19**:252–7.

157. Quillin SP, Brink JA, Heiken JP, Siegel CL, McClennan BL, Clayman RV. Helical (spiral) CT angiography for identification of crossing vessels at the ureteropelvic junction. *AJR Am J Roentgenol* 1996;**166**:1125–30.

158. Quint LE, Francis IR, Williams DM, Bass JC, Shea MJ, Frayer DL, *et al.* Evaluation of thoracic aortic disease with the use of helical CT and multiplanar reconstructions: comparison with surgical findings. *Radiology* 1996;**201**:37–41.

159. Roos YB, Hasan D, Vermeulen M. Outcome in patients with large intraventricular haemorrhages: a volumetric study. *J Neurol Neurosurg Psychiatry* 1995;**58**:622–4.

160. Sagy M, Poustchiamin M, Nimkoff L, Silver P, Shikowitz M, Leonidas JC. Spiral computed tomographic scanning of the chest with 3-dimensional imaging in the diagnosis and management of pediatric intrathoracic airway obstruction. *Thorax* 1996;**51**:1005–9.

161. Tecce PM, Fishman EK. Spiral CT with multiplanar reconstruction in the diagnosis of sternoclavicular osteomyelitis. *Skeletal Radiol* 1995;**24**:275–81.

162. White RA, Donayre CE, Walot I, Kopchok GE, Wilson EP, Buwalda R, *et al.* Preliminary clinical outcome and imaging criterion for endovascular prosthesis development in high-risk patients who have aortoiliac and traumatic arterial lesions. *J Vasc Surg* 1996;**24**:556–69; discussion 569–71.

163. Zimmerman RA, Gusnard DA, Bilaniuk LT. Pediatric craniocervical spiral CT. *Neuroradiology* 1992;**34**:112–16.

164. Anderson WA, Ilkowski DA, Eldredge J, Cha R, Chen C, Waters D, *et al.* The small aortic root and the Medtronic Hall valve: ultrafast computed tomography assessment of left ventricular mass following aortic valve replacement. *J Heart Valve Dis* 1996;**5**(Suppl 3):S329–35.

165. Chareonthaitawee P, Christian TF, Hirose K, Gibbons RJ, Rumberger JA. Relation of initial infarct size to extent of left-ventricular remodeling in the year after acute myocardial infarction. *J Am Coll Cardiol* 1995;**25**:567–73.

166. Kao SC, Kimura K, Smith WL, Sato Y. Tracheomalacia before and after aortosternopexy: dynamic and quantitative assessment by electron beam computed tomography with clinical correlation. *Pediatr Radiol* 1995;**25**(Suppl 1):S187–93.

167. Kimura K, Soper RT, Kao SC, Sato Y, Smith WL, Franken EA. Aortosternopexy for tracheomalacia following repair of esophageal atresia: evaluation by cine-CT and technical refinement. *J Pediatr Surg* 1990;**25**:769–72.

168. Kim T, Murakami T, Oi H, Matsushita M, Kishimoto H, Igarashi H, *et al.* Detection of hypervascular hepatocellular carcinoma by dynamic MRI and dynamic spiral CT. *J Comput Assist Tomogr* 1995;**19**:948–54.

169. Murakami T, Kim T, Oi H, Nakamura H, Igarashi H, Matsushita M, *et al.* Detectability of hypervascular hepatocellular carcinoma by arterial phase images of MR and spiral CT. *Acta Radiol* 1995;**36**:372–6.

170. Dresel S, Stabler A, Scheidler J, Holzknecht N, Tatsch K, Hahn K. Diagnostic approach in acute pulmonary embolism: perfusion scintigraphy versus spiral computed tomography. *Nucl Med Commun* 1995;**16**:1009–15.

171. Sostman HD, Layish DT, Tapson VF, Spritzer CE, DeLong DM, Trotter P, *et al.* Prospective comparison of helical CT and MR imaging in clinically suspected acute pulmonary embolism. *J Magn Reson Imaging* 1996;**6**:275–81.

172. Steiner P, Lund GK, Debatin JF, Steiner D, Nienaber C, Nicolas V, *et al.* Acute pulmonary embolism: value of transthoracic and trans-esophageal echocardiography in comparison with helical CT. *AJR Am J Roentgenol* 1996;**167**:931–6.

173. Woodard PK, Sostman HD, MacFall JR, DeLong DM, McDonald JW, Foo TK, *et al.* Detection of pulmonary embolism: comparison of contrast-enhanced spiral CT and time-of-flight MR techniques. *J Thorac Imaging* 1995;**10**:59–72.

174. Gutfinger DE, Leung CY, Hiro T, Maheswaran B, Nakamura S, Detrano R, *et al.* In-vitro atherosclerotic plaque and calcium quantitation by intravascular ultrasound and electron beam computed tomography. *Am Heart J* 1996;**131**:899–906.

175. Baron RL, Oliver JH, Dodd GD, Nalesnik M, Holbert BL, Carr B. Hepatocellular carcinoma: evaluation with biphasic, contrast-enhanced, helical CT. *Radiology* 1996;**199**:505–11.

176. Frederick MG, McElaney BL, Singer A, Park KS, Paulson EK, McGee SG, *et al.* Timing of parenchymal enhancement on dual-phase dynamic helical CT of the liver: how long does the hepatic arterial phase predominate? *AJR Am J Roentgenol* 1996;**166**:1305–10.

177. Irie T, Takeshita K, Makita K, Yamauchi T, Kusano S. A one-stage method for obtaining CT during arterial portography and hepatic arteriography. *Acta Radiol* 1994;**35**:135–7.

178. Irie T, Kusano S. Contrast-enhanced spiral CT of the liver: effect of injection time on time to peak hepatic enhancement. *J Comput Assist Tomogr* 1996;**20**:633–7.

179. Luker GD, Siegel MJ, Bradley DA, Baty JD. Hepatic spiral CT in children: scan delay time-enhancement analysis. *Pediatr Radiol* 1996;**26**:337–40.

180. Plumley DA, Grosfeld JL, Kopecky KK, Buckwalter KA, Vaughan WG. The role of spiral (helical) computerized tomography with three-dimensional reconstruction in pediatric solid tumors. *J Pediatr Surg* 1995;**30**:317–21.

181. Roche KJ, Genieser NB, Ambrosino MM. Pediatric hepatic CT: an injection protocol. *Pediatr Radiol* 1996;**26**:502–7.

182. Takayasu K, Furukawa H, Wakao F, Muramatsu Y, Abe H, Terauchi T, *et al.* CT diagnosis of early hepatocellular carcinoma: sensitivity, findings, and CT-pathologic correlation. *AJR Am J Roentgenol* 1995;**164**:885–90.

183. Blum AG, Delfau F, Grignon B, Beurrier D, Chabot F, Claudon M, *et al.* Spiral computed tomography versus pulmonary angiography in the diagnosis of acute massive pulmonary embolism. *Am J Cardiol* 1994;**74**:96–8.

184. van Rossum AB, Treurniet FE, Kieft GJ, Smith SJ, Schepers-Bok R. Role of spiral volumetric computed tomographic scanning in the assessment of patients with clinical suspicion of pulmonary embolism and an abnormal ventilation/perfusion lung scan. *Thorax* 1996;**51**:23–8.

185. Detrano R, Tang W, Kang X, Mahaisavariya P, McCrae M, Garner D, *et al.* Accurate coronary calcium phosphate mass measurements from electron beam computed tomograms. *Am J Card Imaging* 1995;**9**:167–73.

186. Detrano R, Hsiai T, Wang SJ, Puentes G, Fallavollita J, Shields P, *et al.* Prognostic value of coronary calcification and angiographic stenoses in patients undergoing coronary angiography. *J Am Coll Cardiol* 1996;**27**:285–90.

187. Fallavollita JA, Kumar K, Brody AS, Bunnell IL, Canty JM. Detection of coronary artery calcium to differentiate patients with early coronary atherosclerosis from luminally normal arteries. *Am J Cardiol* 1996;**78**:1281–5.

188. Goel M, Wong ND, Eisenberg H, Hagar J, Kelly K, Tobis JM. Risk factor correlates of coronary calcium as evaluated by ultrafast computed tomography. *Am J Cardiol* 1992;**70**:977–80.

189. Guerra OR, Janowitz WR, Agatston AS, Mantelle LL, Viamonte M. Coronary artery diameter and coronary risk factors – a study with ultrafast computed tomography. *Am Heart J* 1993;**126**:600–6.

190. Hoeg JM, Feuerstein IM, Tucker EE. Detection and quantitation of calcific atherosclerosis by ultrafast computed tomography in children and young adults with homozygous familial hypercholesterolemia. *Arterioscler Thromb* 1994;**14**:1066–74.

191. Janowitz WR, Agatston AS, Viamonte M. Comparison of serial quantitative evaluation of calcified coronary artery plaque by ultrafast computed tomography in persons with and without obstructive coronary artery disease. *Am J Cardiol* 1991;**68**:1–6.

192. Janowitz WR, Agatston AS, Kaplan G, Viamonte M. Differences in prevalence and extent of coronary artery calcium detected by ultrafast computed tomography in asymptomatic men and women. *Am J Cardiol* 1993;**72**:247–54.

193. Kaufmann RB, Peyser PA, Sheedy PF, Rumberger JA, Schwartz RS. Quantification of coronary artery calcium by electron beam computed tomography for determination of severity of angiographic coronary artery disease in younger patients. *J Am Coll Cardiol* 1995;**25**:626–32.

194. Maher JE, Raz JA, Bielak LF, Sheedy PF, Schwartz RS, Peyser PA. Potential of quantity of coronary artery calcification to identify new risk factors for asymptomatic atherosclerosis. *Am J Epidemiol* 1996;**144**:943–53.

195. Mahoney LT, Burns TL, Stanford W, Thompson BH, Witt JD, Rost CA, *et al.* Coronary risk factors measured in childhood and young adult life are associated with coronary artery calcification in young adults – the muscatine study. *J Am Coll Cardiol* 1996;**27**:277–84.

196. Mautner GC, Mautner SL, Froehlich J, Feuerstein IM, Proschan MA, Roberts WC, *et al.* Coronary artery calcification: assessment with electron beam CT and histomorphometric correlation. *Radiology* 1994;**192**:619–23.

197. Mautner SL, Mautner GC, Froehlich J, Feuerstein IM, Proschan MA, Roberts WC, *et al.* Coronary artery disease: prediction with in vitro electron beam CT. *Radiology* 1994;**192**:625–30.

198. McCollough CH, Kaufmann RB, Cameron BM, Katz DJ, Sheedy PF, Peyser PA. Electron beam CT – use of a calibration phantom to reduce variability in calcium quantitation. *Radiology* 1995;**196**:159–65.

199. Megnien JL, Sene V, Jeannin S, Hernigou A, Plainfosse MC, Merli I, *et al.* Coronary calcification and its relation to extracoronary atherosclerosis in asymptomatic hypercholesterolemic men. The PCV METRA Group. *Circulation* 1992;**85**:1799–807.

200. Megnien JL, Simon A, Lemariey M, Plainfosse MC, Levenson J. Hypertension promotes coronary calcium deposit in asymptomatic men. *Hypertension* 1996;**27**:949–54.

201. Moshage WEL, Achenbach S, Seese B, Bachmann K, Kirchgeorg M. Coronary artery stenoses – 3-dimensional imaging with electrocardiographically triggered, contrast agent-enhanced, electron beam CT. *Radiology* 1995;**196**:707–14.

202. Roig E, Chomka EV, Castaner A, Campo A, Heras M, Rich S, *et al.* Exercise ultrafast computed tomography for the detection of coronary artery disease. *J Am Coll Cardiol* 1989;**13**:1073–81.

203. Rumberger JA, Schwartz RS, Simons B, Sheedy PF, Edwards WD, Fitzpatrick LA. Relation of coronary calcium determined by electron beam computed tomography and lumen narrowing determined by autopsy. *Am J Cardiol* 1994;**73**:1169–73.

204. Rumberger JA, Simons DB, Fitzpatrick LA, Sheedy PF, Schwartz RS. Coronary artery calcium area by electron beam computed tomography and coronary atherosclerotic plaque area – a histopathologic correlative study. *Circulation* 1995;**92**:2157–62.

205. Simon A, Giral P, Levenson J. Extracoronary atherosclerotic plaque at multiple sites and total coronary calcification deposit in asymptomatic men – association with coronary risk profile. *Circulation* 1995;**92**:1414–21.

206. Simons DB, Schwartz RS, Edwards WD, Sheedy PF, Breen JF, Rumberger JA. Noninvasive definition of anatomic coronary artery disease by ultrafast computed tomographic scanning – a quantitative pathological comparison study. *J Am Coll Cardiol* 1992;**20**:1118–26.

207. Wong ND, Kouwabunpat D, Vo AN, Detrano RC, Eisenberg H, Goel M, *et al.* Coronary calcium and atherosclerosis by ultrafast computed tomography in asymptomatic men and women – relation to age and risk factors. *Am Heart J* 1994;**127**:422–30.

208. Craven CM, Naik KS, Blanshard KS, Batchelor AG, Spencer JA. Multispiral three-dimensional computed tomography in the investigation of craniosynostosis: technique optimization. *Br J Radiol* 1995;**68**:724–30.

209. Dula K, Mini R, vanderStelt PF, Lambrecht JT, Schneeberger P, Buser D. Hypothetical mortality risk associated with spiral computed tomography of the maxilla and mandible. *Eur J Oral Sci* 1996;**104**:503–10.

210. Engeler CE, Tashjian JH, Engeler CM, Geise RA, Holm JC, Ritenour ER. Volumetric high-resolution CT in the diagnosis of interstitial lung disease and bronchiectasis: diagnostic accuracy and radiation dose. *AJR Am J Roentgenol* 1994;**163**:31–5.

211. Kasales CJ, Hopper KD, Ariola DN, TenHave TR, Meilstrup JW, Mahraj RP, *et al.* Reconstructed helical CT scans: improvement in *z*-axis resolution compared with overlapped and nonoverlapped conventional CT scans. *AJR Am J Roentgenol* 1995;**164**:1281–4.

212. Luker GD, Lee BC, Erickson KK. Spiral CT of the temporal bone in unsedated pediatric patients. *Am J Neuroradiol* 1993;**14**:1145–50.

213. Moore MM, Cacak RK, Hendee WR. Multisegmented ion chamber for CT scanner dosimetry. *Med Phys* 1981;**8**:640–5.

214. Moore SC, Judy PF, Garnic JD, Kambic GX, Bonk F, Cochran G, *et al.* Prospectively gated cardiac computed tomography. *Med Phys* 1983;**10**:846–55.

215. Nambu K, Suzuki R, Hirakawa K. Cerebral blood flow: measurement with xenon-enhanced dynamic helical CT. *Radiology* 1995;**195**:53–7.

216. Nishizawa K, Iwai K, Matsumoto T, Sakashita K, Iinuma TA, Tateno Y, *et al.* Estimation of the exposure and a risk–benefit analysis for a CT system designed for a lung cancer mass screening unit. *Radiat Protection Dosim* 1996;**67**:101–8.

217. O'Brien JM, Char DH, Tucker N, Gordon KB, Norman D. Efficacy of unanesthetized spiral computed tomography scanning in initial evaluation of childhood leukocoria. *Ophthalmology* 1995; **102**:1345–50.

218. Reynolds MD, Heuscher DJ, Vembar M. Evaluation of spiral CT on a 4th-generation system. *Eur Radiol* 1995;**5**:102–9.

219. Robinson AE, Hill EP, Harpen MD. Radiation dose reduction in pediatric CT. *Pediatr Radiol* 1986;**16**:53–4.

220. Rubin GD, Walker PJ, Dake MD, Napel S, Jeffrey RB, McDonnell CH, *et al.* Three-dimensional spiral computed tomographic angiography: an alternative imaging modality for the abdominal aorta and its branches. *J Vasc Surg* 1993;**18**:656–64; discussion 665.

221. Ruegsegger P, Koller B, Muller R. A micro-tomographic system for the nondestructive evaluation of bone architecture. *Calcif Tissue Int* 1996;**58**:24–9.

222. Suojanen JN, Regan F. Spiral CT scanning of the paranasal sinuses. *Am J Neuroradiol* 1995;**16**:787–9.

223. Tsuchiya K, Makita K, Furui S. 3D-CT angiography of cerebral aneurysms with spiral scanning: comparison with 3D-time-of-flight MR angiography. *Radiat Med* 1994;**12**:161–6.

224. Vade A, Demos TC, Olson MC, Subbaiah P, Turbin RC, Vickery K, *et al.* Evaluation of image quality using 1:1 pitch and 1.5:1 pitch helical CT in children: a comparative study. *Pediatr Radiol* 1996;**26**:891–3.

225. van der Bruggen-Bogaarts BAHA, van der Bruggen HM, van Waes PF, Lammers JW. Assessment of bronchiectasis: comparison of HRCT and spiral volumetric CT. *J Comput Assist Tomogr* 1996;**20**:15–19.

226. Vannucchi L, Simonelli P, Zanfranceschi G, Niccolai F, Piperno GC. Spiral angio-CT investigation of the carotid bifurcation. *Riv Neuroradiol* 1996;**9**(Suppl 2):106–13.

227. Villafana T. Technologic advances in computed tomography. *Curr Opin Radiol* 1991;**3**:275–83.

228. Vock P, Soucek M. Spiral computed tomography in the assessment of focal and diffuse lung disease. *J Thorac Imaging* 1993;**8**:283–90.

229. Wang G, Vannier MW. Longitudinal resolution in volumetric X-ray computerized tomography – analytical comparison between conventional and helical computerized tomography. *Med Phys* 1994;**21**:429–33.

230. Brody AS, Seidel FG, Kuhn JP. CT evaluation of blunt abdominal trauma in children: comparison of ultrafast and conventional CT. *AJR Am J Roentgenol* 1989;**153**:803–6.

231. Zink FE, McCollough CH. The measurement of radiation dose profiles for electron beam computed tomography using film dosimetry. *Med Phys* 1994;**21**:1287–91.

232. Gore SM. Statistics in practice. London: British Medical Association, 1982.

233. Dickersin K, Scherer R, Lefebvre C. Identifying relevant studies for systematic reviews. *BMJ* 1994;**309**:1286–91.

234. Sheldon TA. Problems of using modelling in the economic evaluation of health care. *Health Econ* 1996;**5**:1–11.

235. Buxton MJ, Drummond MF, van Hout BA, Prince RL, Sheldon TA, Szucs T, *et al.* Modelling in economic evaluation: an unavoidable fact of life. *Health Econ* 1997;**6**:217–27.

236. Steiner P, Phillips F, Wesner D, Lund GK, Kreymann G, Nicolas V, *et al.* Primary diagnosis and follow up in acute pulmonary embolism: comparison of digital subtraction angiography and spiral CT. *Rofo Fortschr Geb Rontgenstr Neuen Bildgeb Verfahr* 1994;**161**:285–91.

237. Eggen DA, Strong JP, McGill HC. Coronary calcification: relationship to clinically significant coronary lesions and race, sex and topographic distribution. *Circulation* 1965;**32**:948–55.

238. Mayo JR, Remy-Jardin M, Muller NL, Remy J, Worsley DF, Hossein-Foucher C, *et al.* Pulmonary embolism: prospective comparison of spiral CT with ventilation–perfusion scintigraphy. *Radiology* 1997;**205**:447–52.

239. Ferretti GR, Bosson JL, Buffaz PD, Ayanian D, Pison C, Blanc F, *et al.* Acute pulmonary embolism: role of helical CT in 164 patients with intermediate probability at ventilation–perfusion scintigraphy and normal results at duplex US of the legs. *Radiology* 1997;**205**:453–8.

240. Cross JJ, Kemp PM, Walsh CG, Flower CD, Dixon AK. A randomized trial of spiral CT and ventilation perfusion scintigraphy for the diagnosis of pulmonary embolism. *Clin Radiol* 1998;**53**:177–82.

241. Dixon AK, Coulden RA. Coronary artery calcification on computed tomography [Editorial]. *Lancet* 1997;**350**:1265.

242. Dixon AK. The impact of medical imaging on the physician's diagnostic and therapeutic thinking. *Eur Radiol* 1998;**8**:488–90.

243. Dixon AK. Evidence-based diagnostic radiology. *Lancet* 1997;**350**:509–12.

Authors' publications/ presentations

Berry E, Beckmann EC. The systematic literature review: what it is and how information technology can help. infoRAD™ exhibit at Radiology UK 96 (May 1996, Birmingham).
Abstract: *Br J Radiol* 1996:**69**(Suppl):268. This presentation is now available on the Internet at: http://agora.leeds.ac.uk/comir/people/eberry/sysrev/sysrev.htm

Berry E, Smith MA. Systematic reviews in diagnostic imaging.
Oral presentation at Institute of Physics and Engineering in Medicine Annual Meeting (September 1996, Leeds).

Berry E, Kelly S, Harris KM, Cullingworth J, Gathercole L, Hutton J, *et al.* Development of a methodology for systematic literature reviews in diagnostic imaging. Poster presentation at Radiology 1997: Imaging, Science and Oncology (May 1997, Birmingham). Abstract: *Br J Radiol* 1997;**70**(Suppl):109.

Hutton J, Berry E, Smith MA, Kelly S, Harris KM, Cullingworth J, *et al.* Technology assessment in diagnostic imaging: use of systematic literature reviews.
Oral presentation at 13th International Meeting of the International Society of Technology Assessment in Health Care (May 1997, Barcelona).

Hutton J, Berry E, Smith MA, Kelly S, Harris KM, Cullingworth J, *et al.* Technology assessment in diagnostic imaging: use of systematic literature reviews. Oral presentation at World Congress on Medical Physics and Biomedical Engineering (September 1997, Nice).

Kelly S, Berry E, Roderick P, Harris KM, Cullingworth J, Gathercole L, *et al.* The identification of bias in studies of the diagnostic performance of imaging modalities. *Br J Radiol* 1997;**70**:1028–35.

Hutton J, Kelly S, Berry E, Smith MA. Trial design for economic evaluation of diagnostic technology.
Poster presentation at 14th International Meeting of the International Society of Technology Assessment in Health Care (June 1998, Ottawa).

Kelly S, Berry E, Roderick P, Harris KM, O'Connor PJ, Hutton J, *et al.* Medical imaging: challenges associated with the assessment of study validity in systematic literature reviews.
Oral presentation at Radiology 1998 (June 1998, Birmingham).
Abstract: *Br J Radiol* 1998;**71**(Suppl):72.

O'Connor PJ, Harris KM, Kelly S, Berry E, Hutton J, Cullingworth J, *et al.* Systematic literature review of endoscopic US and gastric cancer.
Poster presentation at Radiology 1998 (June 1998, Birmingham).
Abstract: *Br J Radiol* 1998;**71**(Suppl):98.

Berry E, Kelly S, Hutton J, Harris KM, Roderick P, Boyce JC, *et al.* Systematic literature review: electron beam CT and symptomatic coronary artery disease.
Oral presentation at Radiology 1999 (May 1999, Birmingham).
Abstract: *Br J Radiol* 1999;**72**(Suppl):26.

Appendix 1

Search strategies

Economics searches

MEDLINE

001 exp economics/
002 cost$.tw
003 economic$.tw
004 expens$.tw
005 money$.tw
006 monetary.tw
007 financ$.tw
008 dollar$.tw
009 effectiveness.tw
010 QALY.tw
011 (benefi$ and (impact or management or outcome or utility)).tw
012 (impact and (management or outcome or utility)).tw
013 (management and (outcome or utility)).tw
014 (outcome and utility).tw
015 (exp treatment outcome/ and (benefi$ or impact or management or outcome or utility)).tw
016 1 or 2 or 3 ... or 15

BIDS

001 cost*
002 economic*
003 expens*
004 money*
005 monetary
006 financ*
007 dollar*
008 effectiveness
009 QALY
010 benefi* + (impact, management, outcome, utility)
011 impact + (management, outcome, utility)
012 management + (outcome, utility)
013 outcome + utility
014 1 or 2 or 3 ... 13

Additional search strategy for patient outcome or therapeutic impact studies

MEDLINE

001 exp survival analysis/
002 survival rate/
003 exp prognosis/
004 prognos$.tw
005 surviv$.tw
006 exp "outcome and process assessment (health care)"/
007 health.hw
008 health$.tw
009 outcome.hw
010 outcome$.tw
011 effectiveness.tw
012 efficien$.tw
013 benefi$.tw
014 improve$.tw
015 succe$.tw
016 impact.tw
017 management.tw
018 quality of life.tw
019 exp quality of life/
020 QALY.tw
021 1 or 2 or 3 ... or 20

BIDS

001 surviv*
002 prognos*
003 health*
004 outcome*
005 effectiveness
006 efficien*
007 benefi*
008 improve*
009 succe*
010 impact
011 management
012 quality of life
013 QALY
014 1 or 2 or 3 ... or 13

Diagnostic performance searches

SCT – MEDLINE

001 tomography.hw
002 (compute$ adj3 tomograph$).ti,ab,sh
003 ct.tw
004 sct.tw
005 hct.tw
006 1 or 2 or 3 or 4 or 5
007 spiral.tw
008 helical.tw
009 continuous volume$.tw
010 slip ring.tw
011 double helix.tw
012 (breath adj2 hold$).tw
013 fourth generation.tw
014 7 or 8 or 9 or 10 or 11 or 12 or 13
015 6 and 14
016 svct.tw
017 hvct.tw
018 (hes adj3 (ct or (compute$ adj2 tomograph$))).tw
019 (volumetric$ adj3 (ct or (compute$ adj2 tomograph$))).tw
020 15 or 16 or 17 or 18 or 19

EBCT – MEDLINE

001 tomography.hw
002 (compute$ adj3 tomograph$).ti,ab,sh
003 ct.tw
004 1 or 2 or 3
005 electron beam.tw
006 ultrafast.tw
007 ultra fast.tw
008 ebt.tw
009 fifth generation.tw
010 c100.tw
011 c150.tw
012 5 or 6 or 7 ... or 11
013 4 and 9
014 ebct.tw
015 ufct.tw
016 imatron.tw

017 (cine adj3 (ct or (compute$ adj2 tomograph$))).tw
018 (fast adj3 (ct or (compute$ adj2 tomograph$))).tw
019 10 or 11 or 12 ... or 15

SCT – BIDS

001 compute* tomograph*
002 compute* # tomograph*
003 compute* # # tomograph*
004 (computed-tomograph*)
005 (computer-tomograph*)
006 (computerised-tomograph*)
007 (computerized-tomograph*)
008 ct
009 sct
010 hct
011 1 or 2 or 3 ... or 10
012 spiral
013 helical
014 continuous volume*
015 (continuous-volume*)
016 slip ring
017 (slip-ring)
018 double helix
019 (double-helix)
020 breath hold*
021 breath # hold*
022 hold* # breath
023 (breath-hold*)
024 fourth generation
025 (fourth-generation)
026 hes
027 12 or 13 or 14 ... or 26
028 11 and 27
029 svct
030 hvct
031 volumetric* ct
032 volumetric* # ct
033 volumetric* # # ct
034 volumetric* tomograph*
035 volumetric* # tomograph*
036 volumetric* # # tomograph*
037 volumetric* # # # tomograph*
038 28 or 29 or 30 ... or 37

EBCT – BIDS

001 compute* tomograph*
002 compute* # tomograph*
003 compute* # # tomograph*
004 (computed-tomograph*)
005 (computer-tomograph*)
006 (computerised-tomograph*)
007 (computerized-tomograph*)
008 ct
009 1 or 2 or 3 ... or 8
010 electron beam
011 (electron-beam)
012 ultrafast
013 ultra fast
014 (ultra-fast)
015 ebt
016 fifth generation
017 (fifth-generation)
018 c100
019 c150
020 (c-100)
021 (c-150)
022 10 or 11 or 12 ... or 21
023 9 and 22
024 ebct
025 ufct
026 imatron
027 cine ct
028 cine # ct
029 cine # # ct
030 (cine-ct)
031 cine tomograph*
032 cine # tomograph*
033 cine # # tomograph*
034 cine # # # tomograph*
035 (cine-tomograph*)
036 fast ct
037 fast # ct
038 fast # # ct
039 (fast-ct)
040 fast tomograph*
041 fast # tomograph*
042 fast # # tomograph*
043 fast # # # tomograph*
044 (fast-tomograph*)
045 23 or 24 or 25 ... or 44

SCT – EMBASE

001 computer-assisted-tomography.de
002 spiral
003 helical
004 volumetric
005 1 and (2 or 3 or 4)
006 liver-cancer.de
007 5 and 6

EBCT – EMBASE

001 computer-assisted-tomography.de
002 electron
003 beam
004 2 and 3
005 ultrafast
006 4 or 5
007 1 and 6
008 coronary-artery-disease.de
009 7 and 8

Radiation dose search keywords

Dosimetry
Dose
Radiation
Exposure
Rando
TLD
Thermoluminescent
Ionization

Appendix 2

Economics checklist

A Type of analysis

A1 Analytical perspective

- [ ] Individual patient
- [ ] Specific institution
- [ ] Target group for specific services
- [ ] Ministry of Health budget
- [ ] Government budget
- [ ] Community or society

A2 Type of analysis

- [ ] Cost description
- [ ] Cost outcome description
- [ ] Cost-comparison analysis
- [ ] Cost-effectiveness analysis
- [ ] Cost–benefit analysis
- [ ] Cost–utility analysis
- [ ] Cost-minimisation analysis
- [ ] Other

A3 Was there comparison of two or more alternatives?	Yes	No	?
A4 Were costs of the alternatives examined?	Yes	No	?
A5 Were consequences of the alternatives examined?	Yes	No	?

B Outcome indicator

B1 Type of outcome indicator

- [ ] Intermediate end-points e.g. sensitivity
- [ ] Clinical end-points e.g. impact on survival
- [ ] Patient outcome e.g.
  - [ ] Disease-specific QoL
  - [ ] Generic QoL
  - [ ] Utility
  - [ ] Willingness to pay

B2 Was outcome indicator appropriate for type of analysis?	Yes	No	?

C Cost analysis

C1 Was there a comprehensive range of costs?	Yes	No	?
C2 Were costs measured as opposed to estimated?	Yes	No	?
C3 Were capital costs considered?	Yes	No	?
C4 Were direct and indirect costs separated?	Yes	No	?
C5 Was discounting used?	Yes	No	N/App
C6 Was there a standardised price base?	Yes	No	?

D Sensitivity analysis

D1 Was sensitivity analysis performed?	Yes	No	?
D2 Was it for all variables with an observed distribution of values?	Yes	No	?
D3 Was it for all major assumptions on variables not observed?	Yes	No	?
D4 Was threshold analysis performed?	Yes	No	?

Appendix 3

Diagnostic performance checklists

Bias checklist with specific guidelines

1. Article details

1.1 Title
1.2 Main author
1.3 What question(s) is the paper addressing?
1.4 Are these questions of value to the specific aims? Yes No

Aims: Health economics
Patient outcome
Therapeutic impact
Diagnostic impact
Diagnostic performance

Patient selection biases

A Referral bias

Questions A1–A3 provide only information. A judgement from this information is required to assess the presence or absence of the three referral biases.

A1 Is the establishment(s) where the study was undertaken stated?

[ ] Yes = The establishment(s) is stated in the text or the origin of the establishment(s) is identifiable from the authors' correspondence addresses. The establishment is the place of origin of the study, such as a university hospital or a cancer institute.

[ ] No = It is not stated and it is unclear from which author's establishment the study was conducted.

A2 Is the establishment from where the patients were referred stated?

[ ] Yes = It is clearly stated in the text; for example, referred from local general practices.

[ ] No = It is not stated.

A3 Is the access to the establishment described?

[ ] Yes = It is stated that the establishment is open access, referral based, public or private etc.

[ ] No = No information.

B Patient filtering bias

B1 Are specific eligibility criteria stated for those included/excluded?

[ ] Yes = Criteria are either reported for all those who do receive the test or those who do not, and the total number of patients referred is given as well as the number included/excluded; or it is clear that all patients referred to the centre receive the diagnostic test.

[ ] No = Criteria or numbers are not reported.

[ ] ? = Insufficient information.

B2 Is diagnostic safety bias present or evident in the eligibility criteria?

[ ] Yes = It is clear that selected patients are excluded to avoid the 'unnecessary' diagnostic test for reasons of safety or exposure.

[ ] No = It is clear that all patients are included despite safety.

[ ] ? = Insufficient information.

B3 Is co-intervention bias present?

[ ] Yes = A selected proportion of the study group received additional interventions. Such interventions include any prior surgery, treatment or tests that are likely to influence the final test performance. This is also known as 'treatment paradox'.[41]

[ ] No = It is stated that all or none of the study group received additional interventions.

[ ] ? = Insufficient information.

B4 Is co-intervention bias avoided via the eligibility criteria?
[ ] **Yes** = It is clearly stated that patients are excluded if they have had additional interventions.
[ ] **No** = It is clear that patients were included despite co-interventions.
[ ] **?** = Insufficient information.

C Patient cohort bias

Questions C1–C3 provide only information. A judgement from this information is required to assess the presence or absence of patient cohort bias.

C1 Are the study group's clinical details described?
[ ] **Yes** = Severity or chronicity of symptoms is reported, together with sex ratio, age range and mean age of both the initial referral group and those receiving the gold standard test.
[ ] **No** = Neither severity nor chronicity, or fewer than three of the demographic characters, are reported; or demographics are not given for both groups.

C2 Are the study group's pathological details described?
[ ] **Yes** = Type and location of disease are reported **for those receiving the gold standard**.
[ ] **No** = None or only one of the above is reported.

C3 Are the study group's co-morbid details described?
[ ] **Yes** = **Any** co-morbid conditions, or absence of conditions are reported for **any** patients.
[ ] **No** = No information regarding co-morbid conditions is reported.

Biases associated with application of the gold standard

D1 Is verification bias present?
[ ] **Yes** = Not all of the patients who have received the diagnostic test go on to receive the gold standard.
[ ] **No** = All patients receive the single gold standard test or a correction is performed by the authors.
[ ] **?** = Insufficient information.

D2 Is work-up bias present?
[ ] **Yes** = The result of the diagnostic test is used to decide who receives the gold standard.
[ ] **No** = It is clear that the diagnostic test is not used to decide, or a correction is performed by the authors.
[ ] **?** = Insufficient information.

D3 Is incorporation bias present?
[ ] **Yes** = Patients receive verification via the diagnostic test under evaluation.
[ ] **No** = The diagnostic test is not used as verification.
[ ] **?** = Insufficient information.

Biases due to the measurement of results

E Disease progression bias

E1 Is disease progression bias present for the test under evaluation?
[ ] **Yes** = The time between the diagnostic test and verification with the gold standard is at least *n* days. The number of days, *n*, considered acceptable depends on the aetiology and understanding of the condition under review.
[ ] **No** = The time is less than *n* days.
[ ] **?** = Insufficient information.

F Withdrawal bias

F1 Are results reported for all patients who received verification?
[ ] **Yes** = Results are clearly reported for all patients who received verification with the gold standard test.
[ ] **No** = Results are missing or selected results are reported.
[ ] **?** = Insufficient information.

F2 Are there any indeterminate test results?
[ ] **Yes** = Patients are excluded or results not reported owing to indeterminate test results.
[ ] **No** = All results are included irrespective of indeterminability.
[ ] **?** = Insufficient information.

F3 Are there any patients lost to follow-up?
[ ] **Yes** = Patients are excluded or results not reported owing to loss.
[ ] **No** = All patients present for verification.
[ ] **?** = Insufficient information.

G Observer variability bias

G1 Is there a single observer of the diagnostic test under evaluation?
[ ] **Yes** = All images from the test under evaluation are interpreted by one person.
[ ] **No** = More than one interpreter.
[ ] **?** = Insufficient information.

G2 If 'no' to G1, are results reported separately for each observer?
[ ] **Yes** = All results are reported independently for all observers.
[ ] **No** = Not all results are reported separately (i.e. pooled).
[ ] **?** = Insufficient information.

G3 Is any attempt made to assess interobserver variability?
[ ] **Yes** = Data are reported statistically, with the kappa statistic, or illustrated using ROC curves for interobserver variation.
[ ] **No** = No data are provided.
[ ] **?** = Insufficient information.

G4 Are the diagnostic test results taken from a consensus decision?
[ ] **Yes** = It is clearly stated that the test results are a consensus decision.
[ ] **No** = It is clear that it was not a consensus decision.
[ ] **?** = Insufficient information.

G5 Is any attempt made to assess intraobserver variability?
[ ] **Yes** = Data are reported statistically, with the kappa statistic, or illustrated using ROC curves for intraobserver variation.
[ ] **No** = No data are provided.
[ ] **?** = Insufficient information.

Independence of interpretation biases

H1 Is diagnostic review bias present?
[ ] **Yes** = Observers are aware of the results of the diagnostic test when interpreting the gold standard.
[ ] **No** = It is stated that observers are blinded or unaware of the diagnostic test results.
[ ] **?** = Insufficient information.

H2 Is test review bias present?
[ ] **Yes** = Observers are aware of the results of the gold standard when interpreting the diagnostic test.
[ ] **No** = It is stated that the observers are blinded or unaware of the gold standard results.
[ ] **?** = Insufficient information.

H3 Is comparator review bias present?
[ ] **Yes** = More than one diagnostic test is compared with the gold standard and observers are aware of the result of one test when interpreting the other test.
[ ] **No** = It is stated that all the diagnostic tests were read independently or blind to the other tests; or only one diagnostic test was used.
[ ] **?** = Insufficient information.

H4 Is clinical review bias present?
[ ] **Yes** = It is stated that the observers are aware of the clinical details and history of the patients.
[ ] **No** = It is stated that the observers are blinded to the clinical data.
[ ] **?** = Insufficient information.

Factors checklists

SCT – liver lesions

1. Article details

1.1 Title

1.2 Main author

1.3 Over what period was the study performed? .....

2. Study cohort

A1 Was the study randomised? Yes No ?

A2 Was the study prospective? Yes No ?

A3 Was the study controlled? Yes No ?

3. Sample size

B1 What was the total number of patients referred? .....

B2 How many patients were excluded or lost?

Before receiving test .....

After receiving test .....

B3 How many positive patients were there? .....

B4 How many negative patients were there? .....

B5 How many lesions were there? .....

B6 Were patients divided into subgroups? Yes No ?

1)	11) .....
2)	12) .....
3)	13) .....
4)	14) .....
5)	15) .....
6)	16) .....
7)	17) .....
8)	18) .....
9)	19) .....
10)	20) .....

4. Clinical description

C1 Number male .....

C2 Number female .....

C3 Age range .....

C4 Mean age .....

C5 Weight

C6 **Were symptoms/diagnosis/indications described?** Yes No ?

.....
.....
.....
.....
.....
.....

5. Homogeneity of diagnostic application

D1 **Main diagnostic application(s)** .....
.....

D2 **Diagnostic application(s) subset(s)** .....
.....
.....

D3 **Diagnostic modality(ies)** .....

D4 **Anatomical area(s) subset(s)** .....

D5 **Tumour types** .....
.....

D6 **Multiple tumour cut-off** .....
.....
.....
.....

6. Technical quality

E1 **Model(s)** .....

E2 **Manufacturer(s)** .....

E3 **Protocol** .....
.....
.....
.....

7. Procedural quality

F1 **Suggested operator ability** .....

F2 **Number of readers** .....

F3 **Diagnostic criteria (thresholds/scorings)** .....
.....
.....
.....
.....
.....

F4 **Contrast**
.....
.....
.....

F5 **Gold standards**
.....
.....
.....

SCT – PE

1. Article details

1.1 Title

1.2 Main author

1.3 Over what period was the study performed? .....

2. Study cohort

A1 Was the study randomised? Yes No ?

A2 Was the study prospective? Yes No ?

A3 Was the study controlled? Yes No ?

3. Sample size

B1 What was the total number of patients referred?

B2 How many patients were excluded or lost?

Before receiving test .....

After receiving test .....

B3 How many true-positive patients were there in the verified group? .....

B4 How many true-negative patients were there in the verified group? .....

B5 Were patients divided into subgroups? Yes No ?

1)	11) .....
2)	12) .....
3)	13) .....
4)	14) .....
5)	15) .....
6)	16) .....
7)	17) .....
8)	18) .....
9)	19) .....
10)	20) .....

4. Clinical description

C1 Number male .....

C2 Number female .....

C3 Age range .....

C4 Mean age .....

C5 Were symptoms/diagnosis/indications described? Yes No ?

.....

.....

.....

.....

5. Homogeneity of diagnostic application

D1 Main diagnostic application(s) .....

.....

D2 Diagnostic application(s) subset(s) .....

.....

.....

6. Technical quality

E1 Scanner

Model(s) .....

Manufacturer(s) .....

E2 Injector model(s)/manufacturer(s) .....

E3 Contrast product(s)/manufacturer(s) .....

7. Procedural quality

F1 Scan procedure

Preliminary scan
Anatomical area .....
kVp .....
mA .....
Scanning time .....
Section thickness/rotation rate .....
Table speed .....

Arterial scan
Anatomical area .....
kVp .....
mA .....
Scanning time .....
Section thickness/rotation rate .....
Table speed .....
Scan delay .....
Breath-hold/shallow breathing/unknown .....
Scan direction: craniocaudal/caudocranial .....
Injection procedure .....
Contrast concentration .....
Contrast amount .....
Injection rate .....

F2 Image presentation
Film/monitor .....
Transverse/coronal/sagittal/oblique .....
Spiral reconstruction slices .....
Window settings .....

F3 Analysis
Analysis procedure .....
Suggested observer ability .....
Number of observers .....

F4 Diagnostic criteria (thresholds/scorings) .....
.....
.....
.....
.....
.....
.....

F5 Gold standards .....
.....
.....
.....

EBCT – CAD

1. Article details

1.1 Title

1.2 Main author

1.3 Over what period was the study performed? .....

2. Study cohort

A1 Was the study randomised? Yes No ?

A2 Was the study prospective? Yes No ?

A3 Was the study controlled? Yes No ?

3. Sample size

B1 What was the total number of patients referred? .....

B2 How many patients were excluded or lost?

Before receiving test .....

After receiving test .....

B3 How many true-positive patients were there in the verified group? .....

B4 How many true-negative patients were there in the verified group? .....

B5 Were patients divided into subgroups? Yes No ?

1) ..... 11) .....
2) ..... 12) .....
3) ..... 13) .....
4) ..... 14) .....
5) ..... 15) .....
6) ..... 16) .....
7) ..... 17) .....
8) ..... 18) .....
9) ..... 19) .....
10) ..... 20) .....

4. Clinical description

C1 Number male .....

C2 Number female .....

C3 Age range .....

C4 Mean age .....

C5 Were symptoms/diagnosis/indications described? Yes No ?

.....
.....
.....
.....

5. Homogeneity of diagnostic application

D1 **Main diagnostic application(s)** .....
.....

D2 **Diagnostic application(s) subset(s)** .....
.....
.....

D3 **Diagnostic modality(ies)** .....

D4 **Main anatomical area(s)** .....
.....
.....

D5 **Anatomical area(s) subset(s)** .....
.....
.....

6. Technical quality

E1 **Model(s)** .....

E2 **Manufacturer(s)** .....

E3 **Protocol** .....
.....
.....
.....
.....

7. Procedural quality

F1 **Suggested operator ability** .....

F2 **Number of readers** .....

F3 **Diagnostic criteria (thresholds/scorings)** .....
.....
.....
.....
.....

Appendix 4

Checklist results and raw data from primary studies

Raw data for SCT (PE only) and EBCT (symptomatic CAD) studies are presented in *Tables 49–51*. Results from the completion of the checklists for the EBCT (symptomatic CAD) studies appear in *Tables 52* and *53*.

TABLE 49 *Raw data from studies of central PE detection*

Author	TP	FN	FP	TN	Sensitivity	Specificity	Accuracy	PPV	NPV	OR
Goodman *et al.* 1995[87]	6	1	1	12	0.86	0.92	0.90	0.86	0.92	72.00
Remy-Jardin *et al.* 1992[20]	18	0	1	23	1.00	0.96	0.98	0.95	1.00	N/App
Remy-Jardin *et al.* 1996[69]	39	4	7	25	0.91	0.78	0.85	0.85	0.86	34.82

TABLE 50 *Raw data from studies of central and peripheral PE detection*

Author	TP	FN	FP	TN	Sensitivity	Specificity	Accuracy	PPV	NPV	OR
Goodman *et al.* 1995[87]	7	4	1	8	0.64	0.89	0.75	0.88	0.67	14.00
van Rossum *et al.* 1996[88]	10	5	0	41	0.67	1.00	0.91	1.00	0.89	N/App
van Rossum *et al.* 1996[88]	12	3	0	41	0.80	1.00	0.95	1.00	0.93	N/App

TABLE 51 *EBCT diagnosis of symptomatic CAD: raw data from studies included in the quantitative analysis*

Author	TP	FN	FP	TN	Sensitivity	Specificity	Accuracy	PPV	NPV	OR
Barbir *et al.* 1994[91]	15	3	31	53	0.833	0.631	0.667	0.326	0.946	8.55
Bormann *et al.* 1992[93]	15	1	25	9	0.938	0.265	0.480	0.375	0.900	5.40
Breen *et al.* 1992[95]	47	0	28	25	1.000	0.472	0.720	0.627	1.000	N/App
Budoff *et al.* 1996[96]	404	23	159	124	0.946	0.438	0.744	0.718	0.844	13.70
Fallavollita *et al.* 1994[98]	50	9	26	21	0.847	0.447	0.670	0.658	0.700	4.49
Rumberger *et al.* 1995[99]	64	1	16	30	0.985	0.652	0.847	0.800	0.968	120.0
Yaghoubi *et al.* 1995[101]	32	1	15	19	0.970	0.559	0.761	0.681	0.950	40.53

TABLE 52 *EBCT diagnosis of symptomatic CAD: results from bias checklist for studies included in quantitative analysis; the questions A–H are those presented in the bias checklist shown in appendix 3*

	Barbir *et al.* 1994[91]	**Bormann *et al.* 1992[93]**	**Breen *et al.* 1992[95]**	**Budoff *et al.* 1996[96]**	**Fallavollita *et al.* 1994[98]**	**Rumberger *et al.* 1995[99]**	**Yaghoubi *et al.* 1995[101]**	**Ideal**
A1	No	Yes	Yes	Yes	Yes	Yes	Yes	Yes
A2	No	No	No	No	No	Yes	No	Yes
A3	No	No	No	No	Yes	Yes	No	Yes
B1	Yes	No	Yes	Yes	No	Yes	No	Yes
B2	?	?	Yes	No	?	No	?	No
B3	?	?	?	?	?	No	?	No
B4	?	?	?	?	?	Yes	?	No
C1	No	Yes	Yes	Yes	Yes	Yes	Yes	Yes
C2	No	Yes	Yes	Yes	Yes	No	No	Yes
C3	No	No	Yes	No	Yes	Yes	No	Yes
D1	No	?	No	No	No	No	?	No
D2	No	?	No	No	No	No	No	No
D3	No	No	No	No	No	No	No	No
E1	Yes	No	No	No	No	No	Yes	No
F1	Yes	Yes	Yes	Yes	No	Yes	Yes	Yes
F2	No	No	No	No	Yes	No	No	No
F3	No	No	No	No	Yes	No	No	No
G1	Yes	No	?	No	No	?	Yes	Yes or No
G2	–	No	No	No	No	No	–	Yes
G3	No	No	No	No	Yes	No	Yes	Yes
G4	No	?	?	?	Yes	?	No	No
G5	No	No	No	No	Yes	No	No	Yes
H1	No	?	No	Yes	No	?	?	No
H2	No	?	No	Yes	?	?	?	No
H3	No	No	No	No	No	No	No	No
H4	No	?	?	?	?	?	?	No

TABLE 53 *EBCT diagnosis of symptomatic CAD: results from factor checklist for studies included in quantitative analysis*

	Barbir *et al.* 1994[91]	**Bormann *et al.* 1992**[93]	**Breen *et al.* 1992**[95]	**Budoff *et al.* 1996**[96]	**Fallavollita *et al.* 1994**[98]	**Rumberger *et al.* 1995**[99]	**Yaghoubi *et al.* 1995**[101]
Randomised	No	No	No	No	No	No	No
Prospective	Yes	Yes	Yes	Yes	No	Yes	Yes
Controlled	No	No	No	No	No	No	No
No. patients	102	50	100	710	108	139	67
Positive	41	16	71	427	59	100	33
Negative	61	34	29	283	27	39	34
Male	88	27	91	456	78	89	32
Female	14	23	9	254	28	50	35
Mean age (years)	M: 53.5 F: 51.2	57.9	47.1	56	43.6	M: 47 ± 7 F: 56 ± 11	55
Age range (years)	N/S	37–82	23–59	24–86	25–49	N/S	30–75
Vessels included							
Left main artery	Yes	Yes	N/S	Yes	Yes	N/S	N/S
Left anterior	Yes	Yes	N/S	Yes	Yes	N/S	N/S
Left circumflex	Yes	Yes	N/S	Yes	Yes	N/S	N/S
Right coronary	Yes	Yes	N/S	Yes	Yes	N/S	N/S

Health Technology Assessment panel membership

This report was identified as a priority by the Diagnostics and Imaging Panel.

Acute Sector Panel

Current members

Chair:
Professor Francis H Creed, University of Manchester

Professor Clifford Bailey, University of Leeds

Ms Tracy Bury, Chartered Society of Physiotherapy

Professor Collette Clifford, University of Birmingham

Dr Katherine Darton, M.I.N.D.

Mr John Dunning, Papworth Hospital, Cambridge

Mr Jonathan Earnshaw, Gloucester Royal Hospital

Mr Leonard Fenwick, Freeman Group of Hospitals, Newcastle-upon-Tyne

Professor David Field, Leicester Royal Infirmary

Ms Grace Gibbs, West Middlesex University Hospital NHS Trust

Dr Neville Goodman, Southmead Hospital Services Trust, Bristol

Professor Mark P Haggard, MRC

Professor Robert Hawkins, University of Manchester

Dr Duncan Keeley, General Practitioner, Thame

Dr Rajan Madhok, East Riding Health Authority

Dr John Pounsford, Frenchay Hospital, Bristol

Dr Mark Sculpher, University of York

Dr Iqbal Sram, NHS Executive, North West Region

Past members

Professor John Farndon, University of Bristol*

Professor Senga Bond, University of Newcastle-upon-Tyne

Professor Ian Cameron, Southeast Thames Regional Health Authority

Ms Lynne Clemence, Mid-Kent Health Care Trust

Professor Cam Donaldson, University of Aberdeen

Professor Richard Ellis, St James's University Hospital, Leeds

Mr Ian Hammond, Bedford & Shires Health & Care NHS Trust

Professor Adrian Harris, Churchill Hospital, Oxford

Dr Gwyneth Lewis, Department of Health

Mrs Wilma MacPherson, St Thomas's & Guy's Hospitals, London

Dr Chris McCall, General Practitioner, Dorset

Professor Alan McGregor, St Thomas's Hospital, London

Professor Jon Nicholl, University of Sheffield

Professor John Norman, University of Southampton

Professor Michael Sheppard, Queen Elizabeth Hospital, Birmingham

Professor Gordon Stirrat, St Michael's Hospital, Bristol

Dr William Tarnow-Mordi, University of Dundee

Professor Kenneth Taylor, Hammersmith Hospital, London

Diagnostics and Imaging Panel

Current members

Chair:
Professor Mike Smith, University of Leeds

Dr Philip J Ayres, Leeds Teaching Hospitals NHS Trust

Dr Paul Collinson, Mayday University Hospital, Thornton Heath

Dr Barry Cookson, Public Health Laboratory Service, Colindale

Professor David C Cumberland, University of Sheffield

Professor Adrian Dixon, University of Cambridge

Mr Steve Ebdon-Jackson, Department of Health

Mrs Maggie Fitchett, Association of Cytogeneticists, Oxford

Dr Peter Howlett, Portsmouth Hospitals NHS Trust

Professor Alistair McGuire, City University, London

Dr Andrew Moore, Editor, *Bandolier*

Dr Peter Moore, Science Writer, Ashtead

Professor Chris Price, London Hospital Medical School

Dr William Rosenberg, University of Southampton

Dr Gillian Vivian, Royal Cornwall Hospitals Trust

Dr Greg Warner, General Practitioner, Hampshire

Past members

Professor Michael Maisey, Guy's & St Thomas's Hospitals, London*

Professor Andrew Adam, Guy's, King's & St Thomas's School of Medicine & Dentistry, London

Dr Pat Cooke, RDRD, Trent Regional Health Authority

Ms Julia Davison, St Bartholomew's Hospital, London

Professor MA Ferguson-Smith, University of Cambridge

Dr Mansel Hacney, University of Manchester

Professor Sean Hilton, St George's Hospital Medical School, London

Mr John Hutton, MEDTAP International Inc., London

Professor Donald Jeffries, St Bartholomew's Hospital, London

Dr Ian Reynolds, Nottingham Health Authority

Professor Colin Roberts, University of Wales College of Medicine

Miss Annette Sergeant, Chase Farm Hospital, Enfield

Professor John Stuart, University of Birmingham

Dr Ala Szczepura, University of Warwick

Mr Stephen Thornton, Cambridge & Huntingdon Health Commission

Dr Jo Walsworth-Bell, South Staffordshire Health Authority

* Previous Chair

continued

continued

Methodology Panel

Current members

Chair:
Professor Martin Buxton, Brunel University

Professor Doug Altman, Institute of Health Sciences, Oxford

Dr David Armstrong, Guy's, King's & St Thomas's School of Medicine & Dentistry, London

Professor Nick Black, London School of Hygiene & Tropical Medicine

Professor Ann Bowling, University College London Medical School

Dr Mike Clarke, University of Oxford

Professor Michael Drummond, University of York

Dr Vikki Entwistle, University of Aberdeen

Professor Ewan Ferlie, Imperial College, London

Professor Ray Fitzpatrick, University of Oxford

Professor Jeremy Grimshaw, University of Aberdeen

Dr Stephen Harrison, University of Leeds

Mr John Henderson, Department of Health

Professor Richard Lilford, Regional Director, R&D, West Midlands

Professor Theresa Marteau, Guy's, King's & St Thomas's School of Medicine & Dentistry, London

Dr Henry McQuay, University of Oxford

Dr Nick Payne, University of Sheffield

Professor Margaret Pearson, NHS Executive North West

Professor David Sackett, Centre for Evidence Based Medicine, Oxford

Dr PAG Sandercock, University of Edinburgh

Dr David Spiegelhalter, Institute of Public Health, Cambridge

Professor Joy Townsend, University of Hertfordshire

Past members

Professor Anthony Culyer, University of York *

Professor Michael Baum, Royal Marsden Hospital

Dr Rory Collins, University of Oxford

Professor George Davey-Smith, University of Bristol

Professor Stephen Frankel, University of Bristol

Mr Philip Hewitson, Leeds FHSA

Mr Nick Mays, King's Fund, London

Professor Ian Russell, University of York

Dr Maurice Slevin, St Bartholomew's Hospital, London

Professor Charles Warlow, Western General Hospital, Edinburgh

Pharmaceutical Panel

Current members

Chair:
Professor Tom Walley, University of Liverpool

Dr Felicity Gabbay, Transcrip Ltd

Mr Peter Golightly, Leicester Royal Infirmary

Dr Alastair Gray, Health Economics Research Unit, University of Oxford

Professor Rod Griffiths, NHS Executive West Midlands

Mrs Jeanette Howe, Department of Health

Professor Trevor Jones, ABPI, London

Ms Sally Knight, Lister Hospital, Stevenage

Dr Andrew Mortimore, Southampton & SW Hants Health Authority

Mr Nigel Offen, Essex Rivers Healthcare, Colchester

Mrs Marianne Rigge, The College of Health, London

Mr Simon Robbins, Camden & Islington Health Authority, London

Dr Frances Rotblat, Medicines Control Agency

Dr Eamonn Sheridan, St James's University Hospital, Leeds

Mrs Katrina Simister, Liverpool Health Authority

Dr Ross Taylor, University of Aberdeen

Past members

Professor Michael Rawlins, University of Newcastle-upon-Tyne*

Dr Colin Bradley, University of Birmingham

Professor Alasdair Breckenridge, RDRD, Northwest Regional Health Authority

Ms Christine Clark, Hope Hospital, Salford

Mrs Julie Dent, Ealing, Hammersmith & Hounslow Health Authority, London

Mr Barrie Dowdeswell, Royal Victoria Infirmary, Newcastle-upon-Tyne

Dr Tim Elliott, Department of Health

Dr Desmond Fitzgerald, Mere, Bucklow Hill, Cheshire

Professor Keith Gull, University of Manchester

Dr Keith Jones, Medicines Control Agency

Dr John Posnett, University of York

Dr Tim van Zwanenberg, Northern Regional Health Authority

Dr Kent Woods, RDRD, Trent RO, Sheffield

* Previous Chair

Population Screening Panel

Current members

Chair:
Professor Sir John Grimley Evans,
Radcliffe Infirmary, Oxford

Ms Stella Burnside,
Altnagelvin Hospitals Trust,
Londonderry

Mr John Cairns,
University of Aberdeen

Professor Howard Cuckle,
University of Leeds

Dr Carol Dezateux,
Institute of Child Health,
London

Dr Anne Dixon Brown,
NHS Executive,
Anglia & Oxford

Professor Dian Donnai,
St Mary's Hospital, Manchester

Dr Tom Fahey,
University of Bristol

Mrs Gillian Fletcher,
National Childbirth Trust

Dr JA Muir Gray,
Institute of Health Sciences,
Oxford

Professor Alexander Markham,
St James's University Hospital,
Leeds

Dr Ann McPherson,
General Practitioner, Oxford

Dr Susan Moss,
Institute of Cancer Research

Dr Sarah Stewart-Brown,
University of Oxford

Past members

Dr Sheila Adam,
Department of Health*

Professor George Freeman,
Charing Cross & Westminster
Medical School, London

Dr Mike Gill, Brent & Harrow
Health Authority

Dr Anne Ludbrook,
University of Aberdeen

Professor Theresa Marteau,
Guy's, King's & St Thomas's
School of Medicine
& Dentistry, London

Professor Catherine Peckham,
Institute of Child Health,
London

Dr Connie Smith,
Parkside NHS Trust,
London

Ms Polly Toynbee, Journalist

Professor Nick Wald,
University of London

Professor Ciaran Woodman,
Centre for Cancer
Epidemiology, Manchester

Primary and Community Care Panel

Current members

Chair:
Dr John Tripp,
Royal Devon & Exeter
Healthcare NHS Trust

Mr Kevin Barton,
East London & City
Health Authority

Professor John Bond,
University of Newcastle-
upon-Tyne

Dr John Brazier,
University of Sheffield

Ms Judith Brodie,
Age Concern, London

Mr Shaun Brogan,
Daventry & South Northants
Primary Care Alliance

Mr Joe Corkill,
National Association for
Patient Participation

Dr Nicky Cullum,
University of York

Professor Pam Enderby,
University of Sheffield

Mr Andrew Farmer,
Institute of Health Sciences,
Oxford

Professor Richard Hobbs,
University of Birmingham

Professor Allen Hutchinson,
University of Sheffield

Dr Phillip Leech,
Department of Health

Dr Aidan Macfarlane,
Oxfordshire Health Authority

Professor David Mant,
Institute of Health Sciences,
Oxford

Dr Chris McCall,
General Practitioner, Dorset

Dr Robert Peveler,
University of Southampton

Professor Jennie Popay,
University of Salford

Ms Hilary Scott,
Tower Hamlets Healthcare
NHS Trust, London

Dr Ken Stein,
North & East Devon
Health Authority

Past members

Professor Angela Coulter,
King's Fund, London*

Professor Martin Roland,
University of Manchester*

Dr Simon Allison,
University of Nottingham

Professor Shah Ebrahim,
Royal Free Hospital, London

Ms Cathy Gritzner,
King's Fund, London

Professor Andrew Haines,
RDRD, North Thames
Regional Health Authority

Dr Nicholas Hicks,
Oxfordshire Health Authority

Mr Edward Jones,
Rochdale FHSA

Professor Roger Jones,
Guy's, King's & St Thomas's
School of Medicine
& Dentistry, London

Mr Lionel Joyce,
Chief Executive, Newcastle City
Health NHS Trust

Professor Martin Knapp,
London School of Economics
& Political Science

Professor Karen Luker,
University of Liverpool

Dr Fiona Moss,
Thames Postgraduate Medical
& Dental Education

Professor Dianne Newham,
King's College London

Professor Gillian Parker,
University of Leicester

Dr Mary Renfrew,
University of Oxford

* Previous Chair

continued

National Coordinating Centre for Health Technology Assessment, Advisory Group

Current members

Chair:
Professor John Gabbay,
Wessex Institute for Health Research & Development

Professor Mike Drummond,
Centre for Health Economics, University of York

Ms Lynn Kerridge,
Wessex Institute for Health Research & Development

Dr Ruairidh Milne,
Wessex Institute for Health Research & Development

Ms Kay Pattison,
Research & Development Directorate, NHS Executive

Professor James Raftery,
Health Economics Unit, University of Birmingham

Professor Ian Russell,
Department of Health Sciences & Clinical Evaluation, University of York

Dr Ken Stein,
North & East Devon Health Authority

Professor Andrew Stevens,
Department of Public Health & Epidemiology, University of Birmingham

Past member

Dr Paul Roderick,
Wessex Institute for Health Research & Development

HTA Commissioning Board

Current members

Chair:
Professor Charles Florey,
Department of Epidemiology & Public Health, Ninewells Hospital & Medical School, University of Dundee

Professor Doug Altman,
Director of ICRF/NHS Centre for Statistics in Medicine, Oxford

Professor John Bond,
Professor of Health Services Research, University of Newcastle-upon-Tyne

Mr Peter Bower,
Independent Health Advisor, Newcastle-upon-Tyne

Ms Christine Clark,
Honorary Research Pharmacist, Hope Hospital, Salford

Professor Shah Ebrahim,
Professor of Epidemiology of Ageing, University of Bristol

Professor Martin Eccles,
Professor of Clinical Effectiveness, University of Newcastle-upon-Tyne

Dr Mike Gill,
Director of Public Health & Health Policy, Brent & Harrow Health Authority

Dr Alastair Gray,
Director, Health Economics Research Centre, University of Oxford

Professor Mark Haggard,
MRC Institute of Hearing Research

Dr Jenny Hewison,
Senior Lecturer, Department of Psychology, University of Leeds

Professor Sir Miles Irving
(Programme Director),
Professor of Surgery, University of Manchester, Hope Hospital, Salford

Professor Alison Kitson,
Director, Royal College of Nursing Institute

Dr Donna Lamping,
Senior Lecturer, Department of Public Health, London School of Hygiene & Tropical Medicine

Professor Alan Maynard,
Professor of Economics, University of York

Professor Jon Nicholl,
Director, Medical Care Research Unit, University of Sheffield

Professor Gillian Parker,
Nuffield Professor of Community Care, University of Leicester

Dr Tim Peters,
Reader in Medical Statistics, Department of Social Medicine, University of Bristol

Professor Martin Severs,
Professor in Elderly Health Care, Portsmouth University

Dr Sarah Stewart-Brown,
Director, Institute of Health Sciences, University of Oxford

Professor Ala Szczepura,
Director, Centre for Health Services Studies, University of Warwick

Dr Gillian Vivian,
Consultant, Royal Cornwall Hospitals Trust

Professor Graham Watt,
Department of General Practice, Woodside Health Centre, Glasgow

Professor Kent Woods,
Regional Director of R&D NHS Executive, Trent

Dr Jeremy Wyatt,
Senior Fellow, Health & Public Policy, School of Public Policy, University College, London

Past members

Professor Ian Russell,
Department of Health Sciences & Clinical Evaluation, University of York*

Professor David Cohen,
Professor of Health Economics, University of Glamorgan

Mr Barrie Dowdeswell,
Chief Executive, Royal Victoria Infirmary, Newcastle-upon-Tyne

Dr Michael Horlington,
Head of Corporate Licensing, Smith & Nephew Group Research Centre

Professor Martin Knapp,
Director, Personal Social Services Research Unit, London School of Economics & Political Science

Professor Theresa Marteau,
Director, Psychology & Genetics Research Group, Guy's, King's & St Thomas's School of Medicine & Dentistry, London

Professor Sally McIntyre,
MRC Medical Sociology Unit, Glasgow

Professor David Sackett,
Centre for Evidence Based Medicine, Oxford

Dr David Spiegelhalter,
MRC Biostatistics Unit, Institute of Public Health, Cambridge

Professor David Williams,
Department of Clinical Engineering, University of Liverpool

Dr Mark Williams,
Public Health Physician, Bristol

* Previous Chair

Copies of this report can be obtained from:

The National Coordinating Centre for Health Technology Assessment,
Mailpoint 728, Boldrewood,
University of Southampton,
Southampton, SO16 7PX, UK.
Fax: +44 (0) 1703 595 639 Email: hta@soton.ac.uk
http://www.hta.nhsweb.nhs.uk

ISSN 1366-5278