The measurement and monitoring of surgical adverse events

J Bruce
EM Russell
J Mollison
ZH Krukowski
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  Email: orders@hta.ac.uk
 c/o Direct Mail Works Ltd  Tel: 02392 492 000
 4 Oakwood Business Centre  Fax: 02392 478 555
 Downley, HAVANT PO9 2NP, UK  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.
The measurement and monitoring of surgical adverse events

J Bruce¹*
EM Russell¹
J Mollison¹
ZH Krukowski²

¹ Department of Public Health, University of Aberdeen, UK
² Grampian University Hospitals NHS Trust, Aberdeen, UK

* Corresponding author. Now at Dugald Baird Centre, University of Aberdeen

Competing interests: none declared

Published July 2001

This report should be referenced as follows:


Health Technology Assessment is indexed in Index Medicus/MEDLINE and Excerpta Medical EMBASE. Copies of the Executive Summaries are available from the NCCHTA website (see opposite).
NHS R&D HTA Programme

The NHS R&D Health Technology Assessment (HTA) Programme was set up in 1993 to ensure that high-quality research information on the costs, effectiveness and broader impact of health technologies is produced in the most efficient way for those who use, manage and provide care in the NHS.

Initially, six HTA panels (pharmaceuticals, acute sector, primary and community care, diagnostics and imaging, population screening, methodology) helped to set the research priorities for the HTA Programme. However, during the past few years there have been a number of changes in and around NHS R&D, such as the establishment of the National Institute for Clinical Excellence (NICE) and the creation of three new research programmes: Service Delivery and Organisation (SDO); New and Emerging Applications of Technology (NEAT); and the Methodology Programme.

Although the National Coordinating Centre for Health Technology Assessment (NCCHTA) commissions research on behalf of the Methodology Programme, it is the Methodology Group that now considers and advises the Methodology Programme Director on the best research projects to pursue.

The research reported in this monograph was funded as project number 97/16/04.

The views expressed in this publication are those of the authors and not necessarily those of the Methodology Programme, HTA Programme or the Department of Health. The editors wish to emphasise that funding and publication of this research by the NHS should not be taken as implicit support for any recommendations made by the authors.

Criteria for inclusion in the HTA monograph series
Reports are published in the HTA monograph series if (1) they have resulted from work commissioned for the HTA Programme, and (2) they are of a sufficiently high scientific quality as assessed by the referees and editors.

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

Methodology Programme Director: Professor Richard Lilford
HTA Programme Director: Professor Kent Woods
Series Editors: Professor Andrew Stevens, Dr Ken Stein, Professor John Gabbay and Dr Ruairidh Milne
Monograph Editorial Manager: Melanie Corris

The editors and publisher have tried to ensure the accuracy of this report but do not accept liability for damages or losses arising from material published in this report. They would like to thank the referees for their constructive comments on the draft document.

ISSN 1366-5278
© Queen’s Printer and Controller of HMSO 2001

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

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

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

<table>
<thead>
<tr>
<th>Section</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>List of abbreviations</td>
<td>i</td>
</tr>
<tr>
<td>Executive summary</td>
<td>iii</td>
</tr>
<tr>
<td>1 Introduction</td>
<td>1</td>
</tr>
<tr>
<td>2 Methods</td>
<td>3</td>
</tr>
<tr>
<td>Approach to the review</td>
<td>3</td>
</tr>
<tr>
<td>National surgical data</td>
<td>4</td>
</tr>
<tr>
<td>Remit and coverage of the review</td>
<td>7</td>
</tr>
<tr>
<td>Assessment criteria for definition and measurement</td>
<td>7</td>
</tr>
<tr>
<td>Assessment criteria for monitoring systems</td>
<td>8</td>
</tr>
<tr>
<td>Literature search</td>
<td>9</td>
</tr>
<tr>
<td>3 Definition and measurement of surgical wound infection</td>
<td>13</td>
</tr>
<tr>
<td>Definition</td>
<td>13</td>
</tr>
<tr>
<td>Comparison between standardised definitions</td>
<td>17</td>
</tr>
<tr>
<td>Definitions identified from prospective studies</td>
<td>18</td>
</tr>
<tr>
<td>Grading of surgical wound infection</td>
<td>20</td>
</tr>
<tr>
<td>4 Surveillance and monitoring of surgical wound infection</td>
<td>29</td>
</tr>
<tr>
<td>Quantification of risk</td>
<td>29</td>
</tr>
<tr>
<td>Hospital-based nosocomial surveillance systems</td>
<td>31</td>
</tr>
<tr>
<td>Postdischarge surveillance</td>
<td>34</td>
</tr>
<tr>
<td>Validation literature</td>
<td>38</td>
</tr>
<tr>
<td>5 Critical assessment of surgical wound infection</td>
<td>43</td>
</tr>
<tr>
<td>The definition of surgical wound infection</td>
<td>43</td>
</tr>
<tr>
<td>Formal validation of the definition of surgical wound infection</td>
<td>43</td>
</tr>
<tr>
<td>Formal validation of wound grading scales</td>
<td>46</td>
</tr>
<tr>
<td>Monitoring of surgical wound infection</td>
<td>48</td>
</tr>
<tr>
<td>Discussion</td>
<td>52</td>
</tr>
<tr>
<td>6 Definition and measurement of anastomotic leak</td>
<td>55</td>
</tr>
<tr>
<td>Excluded studies</td>
<td>55</td>
</tr>
<tr>
<td>Studies of oesophagogastric surgery</td>
<td>55</td>
</tr>
<tr>
<td>Studies of hepatopancreaticobiliary surgery</td>
<td>58</td>
</tr>
<tr>
<td>Studies of lower gastrointestinal surgery</td>
<td>60</td>
</tr>
<tr>
<td>Studies of methods and accuracy of detection of anastomotic leak</td>
<td>64</td>
</tr>
<tr>
<td>Critical assessment of the properties of anastomotic leak</td>
<td>69</td>
</tr>
<tr>
<td>Conclusions</td>
<td>72</td>
</tr>
<tr>
<td>7 Definition and measurement of DVT</td>
<td>73</td>
</tr>
<tr>
<td>Diagnostic techniques</td>
<td>73</td>
</tr>
<tr>
<td>Factors for critical appraisal</td>
<td>76</td>
</tr>
<tr>
<td>8 Surgical mortality monitoring systems</td>
<td>77</td>
</tr>
<tr>
<td>Surgical mortality monitoring systems within general surgery</td>
<td>77</td>
</tr>
<tr>
<td>Monitoring systems within other surgical specialties</td>
<td>81</td>
</tr>
<tr>
<td>Generic UK national monitoring systems</td>
<td>89</td>
</tr>
<tr>
<td>Generic US national and regional monitoring systems</td>
<td>90</td>
</tr>
<tr>
<td>9 Discussion</td>
<td>93</td>
</tr>
<tr>
<td>Definition and ascertainment of surgical mortality</td>
<td>93</td>
</tr>
<tr>
<td>Risk adjustment</td>
<td>97</td>
</tr>
<tr>
<td>Denominator data</td>
<td>97</td>
</tr>
<tr>
<td>Morbidity</td>
<td>98</td>
</tr>
<tr>
<td>Data processing and validation</td>
<td>98</td>
</tr>
<tr>
<td>Costs</td>
<td>99</td>
</tr>
<tr>
<td>Output and feedback</td>
<td>99</td>
</tr>
<tr>
<td>Generic versus dedicated monitoring systems</td>
<td>100</td>
</tr>
<tr>
<td>Summary</td>
<td>102</td>
</tr>
<tr>
<td>10 Conclusions</td>
<td>103</td>
</tr>
<tr>
<td>11 Implications and recommendations</td>
<td>107</td>
</tr>
<tr>
<td>Surgical wound infection</td>
<td>107</td>
</tr>
<tr>
<td>Anastomotic leak</td>
<td>107</td>
</tr>
<tr>
<td>Deep vein thrombosis</td>
<td>107</td>
</tr>
<tr>
<td>Monitoring systems</td>
<td>107</td>
</tr>
<tr>
<td>Acknowledgements</td>
<td>109</td>
</tr>
<tr>
<td>References</td>
<td>111</td>
</tr>
<tr>
<td>Appendix 1 Databases searched and search strategies</td>
<td>129</td>
</tr>
<tr>
<td>Appendix 2 Statistical assessment criteria and data extraction forms</td>
<td>137</td>
</tr>
<tr>
<td>Appendix 3 Summary of included studies</td>
<td>141</td>
</tr>
<tr>
<td>Health Technology Assessment reports published to date</td>
<td>187</td>
</tr>
<tr>
<td>Methodology Group</td>
<td>193</td>
</tr>
<tr>
<td>HTA Commissioning Board</td>
<td>194</td>
</tr>
</tbody>
</table>
# List of abbreviations

<table>
<thead>
<tr>
<th>Abbreviation</th>
<th>Definition</th>
</tr>
</thead>
<tbody>
<tr>
<td>ASA</td>
<td>American Society of Anesthesiology</td>
</tr>
<tr>
<td>CABG</td>
<td>coronary artery bypass graft</td>
</tr>
<tr>
<td>CDC</td>
<td>Centers for Disease Control</td>
</tr>
<tr>
<td>CEPOD</td>
<td>Confidential Enquiry into Peri-Operative Deaths</td>
</tr>
<tr>
<td>CI</td>
<td>confidence interval*</td>
</tr>
<tr>
<td>CN</td>
<td>circulating nurse*</td>
</tr>
<tr>
<td>COPPISH</td>
<td>Core Patient Profile Information in Scottish Hospitals</td>
</tr>
<tr>
<td>CT</td>
<td>computed tomography</td>
</tr>
<tr>
<td>DSWI</td>
<td>deep sternal wound infection*</td>
</tr>
<tr>
<td>DVT</td>
<td>deep vein thrombosis</td>
</tr>
<tr>
<td>EDINA</td>
<td>Edinburgh Data &amp; Information Access</td>
</tr>
<tr>
<td>ELFA</td>
<td>enzyme-linked fluorescent immunoassay</td>
</tr>
<tr>
<td>ELISA</td>
<td>enzyme-linked immunosorbent assay</td>
</tr>
<tr>
<td>FCIC</td>
<td>Florida Consortium for Infection Control*</td>
</tr>
<tr>
<td>FINNVASC</td>
<td>Finnish Vascular Surgery Registry</td>
</tr>
<tr>
<td>GRASS</td>
<td>gradient-recalled acquisition in steady state</td>
</tr>
<tr>
<td>HCFA</td>
<td>Health Care Finance Administration</td>
</tr>
<tr>
<td>HMPS</td>
<td>Harvard Medical Practice Study</td>
</tr>
<tr>
<td>IA</td>
<td>instant assay</td>
</tr>
<tr>
<td>ICD</td>
<td>International Classification of Diseases</td>
</tr>
<tr>
<td>ICP</td>
<td>infection control practitioner</td>
</tr>
<tr>
<td>LMR</td>
<td>Landelijke Medische Registratie</td>
</tr>
<tr>
<td>NA</td>
<td>nurse anaesthetist*</td>
</tr>
<tr>
<td>NCEPOD</td>
<td>National Confidential Enquiry into Peri-Operative Deaths</td>
</tr>
<tr>
<td>NINSS</td>
<td>Nosocomial Infection National Surveillance Scheme</td>
</tr>
<tr>
<td>NNIS</td>
<td>National Nosocomial Infections Surveillance</td>
</tr>
<tr>
<td>NPS</td>
<td>National Prevalence Survey</td>
</tr>
<tr>
<td>NPV</td>
<td>negative predictive value*</td>
</tr>
<tr>
<td>NSIH</td>
<td>National Programme for Surveillance of Hospital Infections</td>
</tr>
<tr>
<td>OPCS</td>
<td>Office of Population Censuses and Surveys</td>
</tr>
<tr>
<td>OSI</td>
<td>operative site infection*</td>
</tr>
<tr>
<td>PEDW</td>
<td>Patient Episode Database for Wales</td>
</tr>
<tr>
<td>PHLS</td>
<td>Public Health Laboratory Service</td>
</tr>
<tr>
<td>PO</td>
<td>physician observer*</td>
</tr>
<tr>
<td>PPV</td>
<td>positive predictive value*</td>
</tr>
<tr>
<td>QAHCS</td>
<td>Quality in Australian Health Care Study</td>
</tr>
<tr>
<td>RCT</td>
<td>randomised controlled trial*</td>
</tr>
<tr>
<td>SASM</td>
<td>Scottish Audit of Surgical Mortality</td>
</tr>
<tr>
<td>SCTS</td>
<td>Society of Cardiothoracic Surgeons of Great Britain and Ireland</td>
</tr>
<tr>
<td>SENIC</td>
<td>Study of the Efficacy of Nosocomial Infection Control</td>
</tr>
<tr>
<td>SERNIP</td>
<td>Safety and Efficacy of New Interventional Procedures</td>
</tr>
<tr>
<td>SHEA</td>
<td>Society for Hospital Epidemiology of America</td>
</tr>
</tbody>
</table>

*continued*
### List of abbreviations

**continued**

<table>
<thead>
<tr>
<th>Abbreviation</th>
<th>Full Form</th>
<th>Abbreviation</th>
<th>Full Form</th>
</tr>
</thead>
<tbody>
<tr>
<td>SHFA</td>
<td>Scottish Hip Fracture Audit</td>
<td>SWAS</td>
<td>Southampton Wound Assessment Scale*</td>
</tr>
<tr>
<td>SIRS</td>
<td>systemic inflammatory response syndrome</td>
<td>SWEDVASC</td>
<td>Swedish Vascular Registry</td>
</tr>
<tr>
<td>SISG</td>
<td>Surgical Infections Society Group</td>
<td>SWI</td>
<td>surgical wound infection*</td>
</tr>
<tr>
<td>SMR</td>
<td>Scottish Morbidity Record</td>
<td>TRISS</td>
<td>Trauma Score and Injury Severity Score</td>
</tr>
<tr>
<td>SPECT</td>
<td>single photon emission computed tomography*</td>
<td>UCDSS</td>
<td>Uniform Clinical Data Set System</td>
</tr>
<tr>
<td>SSI</td>
<td>surgical site infection*</td>
<td>UKTARN</td>
<td>UK Trauma Audit and Research Network</td>
</tr>
<tr>
<td>STAG</td>
<td>Scottish Trauma Audit Group</td>
<td></td>
<td></td>
</tr>
<tr>
<td>STS</td>
<td>Society of Thoracic Surgeons</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

* Used only in tables
Background

Surgical adverse events contribute significantly to postoperative morbidity, yet the measurement and monitoring of events is often imprecise and of uncertain validity. Given the trend of decreasing length of hospital stay and the increase in use of innovative surgical techniques – particularly minimally invasive and endoscopic procedures – accurate measurement and monitoring of adverse events is crucial.

Objectives

The aim of this methodological review was to identify a selection of common and potentially avoidable surgical adverse events and to assess whether they could be reliably and validly measured, to review methods for monitoring their occurrence and to identify examples of effective monitoring systems for selected events. This review is a comprehensive attempt to examine the quality of the definition, measurement, reporting and monitoring of selected events that are known to cause significant postoperative morbidity and mortality.

Methods

Selection of surgical adverse events

Four adverse events were selected on the basis of their frequency of occurrence and likelihood of evidence of measurement and monitoring:

- surgical wound infection
- anastomotic leak
- deep vein thrombosis (DVT)
- surgical mortality.

Surgical wound infection and DVT are common events that cause significant postoperative morbidity. Anastomotic leak is a less common event, but risk of fatality is associated with delay in recognition, detection and investigation. Surgical mortality was selected because of the effort known to have been invested in developing systems for monitoring surgical death, both in the UK and internationally. Systems for monitoring surgical wound infection were also included in the review.

Literature search

Thirty separate, systematic literature searches of core health and biomedical bibliographic databases (MEDLINE, EMBASE, CINAHL, HealthSTAR and the Cochrane Library) were conducted. The reference lists of retrieved articles were reviewed to locate additional articles. A matrix was developed whereby different literature and study designs were reviewed for each of the surgical adverse events. Each article eligible for inclusion was independently reviewed by two assessors.

Critical appraisal

Studies were appraised according to predetermined assessment criteria. Definitions and grading scales were assessed for: content, criterion and construct validity; repeatability; reproducibility; and practicality (surgical wound infection and anastomotic leak). Monitoring systems for surgical wound infection and surgical mortality were assessed on the following criteria:

- coverage of the system
- whether or not denominator data were collected
- whether standard and agreed definitions were used
- inclusion of risk adjustment
- issues related to data collection
- postdischarge surveillance
- output in terms of feedback and wider dissemination.

Results

Surgical wound infection

A total of 41 different definitions and 13 grading scales of surgical wound infection were identified from 82 studies. Definitions of surgical wound infection varied from ‘presence of pus’ to complex definitions such as those proposed by the Centres for Disease Control in the USA. A small body of literature has been published on the content, criterion and construct validity of different definitions, and comparisons have been made against wound assessment scales and multidimensional indices. There are examples of comprehensive hospital-
based monitoring systems of surgical wound infection, mainly under the auspices of nosocomial surveillance. To date, however, there is little evidence of systematic measurement and monitoring of surgical wound infection after hospital discharge.

**Anastomotic leak**
Over 40 definitions of anastomotic leak were extracted from 107 studies of upper gastrointestinal, hepatopancreatobiliary and lower gastrointestinal surgery. No formal evaluations were found that assessed the validity or reliability of definitions or severity scales of anastomotic leak. One definition was proposed during a national consensus workshop, but no evidence of its use was found in the surgical literature. The lack of a single definition or gold standard hampers comparison of postoperative anastomotic leak rates between studies and institutions.

**Deep vein thrombosis**
Although a critical review of the DVT literature could not be completed within the realms of this review, it was evident that a number of new techniques for the detection and diagnosis of DVT have emerged in the last 20 years. The group recommends a separate review be undertaken of the different diagnostic tests to detect DVT.

**Surgical mortality monitoring systems**
The definition of surgical mortality is relatively consistent between monitoring systems, but duration of follow-up of death postdischarge varies considerably. The majority of systems report in-hospital mortality rates; only some have the potential to link deaths to national death registers. Risk assessment is an important factor and there should be a distinction between recording pre-intervention factors and postoperative complications. A variety of risk scoring systems was identified in the review. Factors associated with accurate and complete data collection include the employment of local, dedicated personnel, simple and structured prompts to ensure that clinical input is complete, and accurate and automated data capture and transfer.

**Conclusions**
The use of standardised, valid and reliable definitions is fundamental to the accurate measurement and monitoring of surgical adverse events. This review found inconsistency in the quality of reporting of postoperative adverse events, limiting accurate comparison of rates over time and between institutions. The duration of follow-up for individual events will vary according to their natural history and epidemiology. Although risk-adjusted aggregated rates can act as screening or warning systems for adverse events, attribution of whether events are avoidable or preventable will invariably require further investigation at the level of the individual, unit or department.

**Recommendations for research**
- A single, standard definition of surgical wound infection is needed so that comparisons over time and between departments and institutions are valid, accurate and useful. Surgeons and other healthcare professionals should consider adopting the 1992 Centers for Disease Control (CDC) definition for superficial incisional, deep incisional and organ/space surgical site infection for hospital monitoring programmes and surgical audits. There is a need for further methodological research into the performance of the CDC definition in the UK setting.
- There is a need to formally assess the reliability of self-diagnosis of surgical wound infection by patients.
- There is a need to assess formally the reliability of case ascertainment by infection control staff.
- Work is needed to create and agree a standard, valid and reliable definition of anastomotic leak which is acceptable to surgeons.
- A systematic review is needed of the different diagnostic tests for the diagnosis of DVT.
- The following variables should be considered in any future DVT review: anatomical region (lower limb, upper limb, pelvis); patient presentation (symptomatic, asymptomatic); outcome of diagnostic test (successfully completed, inconclusive, technically inadequate, negative); length of follow-up; cost of test; whether or not serial screening was conducted; and recording of laboratory cut-off values for fibrinogen equivalent units.
- A critical review is needed of the surgical risk scoring used in monitoring systems.
- In the absence of automated linkage there is a need to explore the benefits and costs of monitoring in primary care.
- The growing potential for automated linkage of data from different sources (including primary care, the private sector and death registers) needs to be explored as a means of improving the ascertainment of surgical complications, including death. This linkage needs to be within the terms of data protection, privacy and human rights legislation.
- A review is needed of the extent of the use and efficiency of routine hospital data versus special collections or voluntary reporting.
The term ‘surgical adverse event’ is relatively new, but the concept of monitoring surgical outcomes, including postoperative complications, began long ago. Alderson, in an influential paper on the evaluation of health information systems, observed that systems for collecting information about hospital patients were advocated in the *Lancet* in 1732. Processing of vital statistics (births, marriages, deaths) has existed in Great Britain since 1838, and the current national hospital morbidity schemes came with the NHS in 1948. The first focus on adverse events is credited to Codman, who from 1910 campaigned publicly for looking at ‘End Results’; in this he recorded not only end results or outcomes but also reasons for their not being achieved. Awareness of iatrogenic disease came in the 1960s, and challenged the traditional view and acceptance of the professional medical expert. For the next two decades the emphasis was on healthcare evaluation and outcome measurement, and it was not until the Harvard Medical Practice Study (HMPS), published in the early 1990s, that the language of adverse events became common.

The HMPS was the first major project to quantify medical injury empirically, based on a review of over 30,000 medical records from 51 hospitals in New York State. A two-stage review process was used whereby, after initial screening and selection, records were assessed by physicians for adverse events and negligence. An adverse event was defined as “an injury that was caused by medical management (rather than the underlying disease) and that prolonged the hospitalisation, produced a disability at time of discharge, or both”. The estimated state-wide incidence rate of adverse events was 3.7%, and 27.6% of the adverse events were attributed to negligence. Surgical specialties had higher rates of adverse events but not higher rates of negligence, although of all the adverse events identified half resulted from surgery. Operative complications were subclassified as: technical, non-technical, related to wound infections, surgical failure and late complications. Wound infections were the most common surgical adverse event. Many of the identified events, such as adhesive intestinal obstructions, were not preventable, but other unpreventable events occurred with predictable frequency; however patients accepted the risk of treatment because of potential benefits.

In 1995, the Quality in Australian Health Care Study (QAHCS), based on the HMPS study, reported that 16.6% of admissions to Australian hospitals were associated with an adverse event that resulted in a disability or longer hospital stay, caused by healthcare management. Half of the identified adverse events were considered preventable. Both these studies have been widely quoted and have been influential in emphasising the scale of the problem. However, they are single cross-sectional surveys and did not attempt to address the need for longitudinal monitoring and, ideally, early detection.

The expansion of interest in the field of medical injury or error continued throughout the 1990s, and the terminology has evolved to include: medical uncertainty, adverse events, maloccurrences, therapeutic misadventure, iatrogenic injury, mshaps, preventable deaths, errors, negligence and malpractice. Although this interest originated in the USA, partly as a consequence of litigation, events in the UK, such as excessive cardiac deaths in Bristol, fuelled the debate, and the term ‘medical error’ is now frequently used in the medical and national press.

All these shifts of emphasis have resulted in parallel shifts in the focus of research, evaluation and the mainstream of publications. The initial studies published in the early 1990s were an attempt to identify and estimate the frequency and nature of adverse events within institutions. More recently, there has been a general acceptance of the scale of the problem and the emphasis is moving towards how best to tackle it. There is a growing concordance that adverse events should be approached at the level of the system or organisation, rather than at an individual level. Leape and co-workers reported that a “complicated, highly technical system of medical care was provided not only by a diverse group of doctors, other care givers, and support personnel, but also by a medical–industrial system that supplied drugs and equipment, and assessment of events required an examination of all these factors as well as of their relation with each other”.

**Chapter I**

**Introduction**
In 2000, the Department of Health published *An Organisation with a Memory*, a report which highlighted the lack of a single comprehensive approach to the monitoring of adverse healthcare events. An adverse event was defined as “An event or omission during clinical care and causing physical or psychological injury to a patient”. However, as the title suggests, the focus of the report was on the NHS as an organisation rather than solely on clinical interventions. Indeed, it recommended the shift away from the clinician or healthcare professional as the source of error and encouraged examination of organisational processes and possible reasons and factors contributing to breakdown within the system. With respect to mechanisms for identifying and learning from adverse healthcare events, voluntary adverse event reporting systems do exist for medicines (Yellow Care scheme) and devices (Medical Devices Agency), but no such schemes exist for surgery. Voluntary reporting of the Safety and Efficacy of New Interventional Procedures in the UK (SERNIP) was set up by the Academy of Medical Royal Colleges in 1996, but is designed for the early evaluation of new procedures rather than for continuous monitoring.

In a recent review of the adverse event literature, three commonalities were identified:

- adverse events are untoward, undesirable or detrimental
- adverse events have an impact on the patient
- the cause of adverse events is a healthcare process (e.g. omission, commission) rather than the natural process of disease.

Walshe summarised the deficiencies of systems for retrospective identification of adverse events. Most commonly, these are case-note reviews or self-reporting, both of which depend on completeness of recording for the events of interest and of the whole process of patient management. These often are not available unless there has been direct data collection in a sophisticated automated patient monitoring system. Obviously, if no monitoring system exists, then crisis self-reporting can be introduced to increase awareness and act as the basis of an investigation. However, Walshe noted that this approach, first, gives rise to a very negative attitude to quality monitoring, and secondly will contain a wide range of interpretations that may vary according to the context of the suspected event.

The Health Technology Assessment (HTA) brief for this review recognised that self-reporting initiatives such as SERNIP will only be of help if standard, validated definitions and measures of adverse events and surgical complications (e.g. wound infections) are available and used. The review was therefore commissioned to explore the methodology of how to measure and monitor adverse events specific to surgery, based on a review of the current literature. Despite the fact that much of the current emphasis is on acting to reduce ‘one-off’ errors in health organisations, ongoing monitoring of adverse events at the core of the clinical service is vital, particularly if it is to underpin clinical improvement. Within surgery, as within other specialities of medicine, measurement of postoperative events is often imprecise and of uncertain validity. An epidemiological or statistical framework for monitoring, based on sound methodology, permits comparison of the frequency of events in different settings and patients; and it offers the potential to highlight deviation or aberration from accepted background norms, and a basis for setting targets for improvement. The aim of this methodological review was, therefore, to identify best practice in measurement and monitoring of surgical adverse events. The review did not attempt to elucidate the causes of adverse events and, in particular, whether they arise from clinical or systems errors. The adverse events reviewed here were chosen because they were identified as causing significant postoperative morbidity and mortality in the UK. This means that there may exist a model system for monitoring rare surgical adverse events that we have not included in our review. Because the focus was methodological, the literature was very different from what would have been found in a search for the avoidability of events or how they should be investigated once detected. Finally, we have used the term ‘surgical adverse event’ interchangeably with ‘surgical complication’, although the point at which a complication becomes an adverse event is a subject for debate. We believe that it is not possible to have a single definition of what constitutes an ‘adverse event’, certainly in surgery. ‘Adverse’ implies avoidability, but in surgery a complication may be unavoidable one year but avoidable the next because procedures have improved. Our interpretation and definition of a surgical adverse event or complication is an event that is attributable to the operative procedure and can occur at any time postoperatively as long as its occurrence is still attributable to a preceding surgical intervention.
Chapter 2

Methods

This chapter outlines the methods used in the review. It includes: the process of selection of surgical adverse events for inclusion; a description of our interpretation of definition, measurement and monitoring; and expansion upon the methods and sources used to identify and evaluate relevant literature.

Approach to the review

The first stage was to select surgical adverse events or surgical procedures for inclusion in the review. Clearly it was not possible to cover the whole field of adverse events arising from surgical procedures, but it was desirable, on the basis of the whatever conditions were reviewed, to reach some generalisable conclusions about the definition, measurement and monitoring of surgical adverse events as a whole. Preliminary searches on MEDLINE for published literature on postoperative complications generated over 30,000 citations for a 4-year period, thus it was necessary to focus the scope of the review.

The review was approachable in one of two ways: by starting from the surgical procedure or from the surgical adverse event. By using the former, the review group would select a surgical procedure and search for evidence of valid ascertainment of surgical adverse events or complications known to be attributable to the antecedent procedure, thus effectively using an outcomes-based approach, with attribution and risk as the main focus. For example, laparoscopic cholecystectomy could be selected and a search conducted for either all or selected adverse events associated with this procedure (e.g. bile duct injury; bowel rupture). The advantage of this approach would be a comprehensive review of the reporting, and quality of reporting, of adverse events associated with one or more operative procedures.

The second approach was to start from the adverse event rather than the surgery, so that the prime focus was on the event itself rather than its operative aetiology. A good definition would be of little use if it could not be reliably monitored, and therefore there was merit in a match between the choice of events to be reviewed and the ability to monitor them. This approach was much more likely to allow some generalisation and was much closer to the original brief for the review. The review panel therefore opted for the second approach. A review focused on adverse events/surgical complications and how they are monitored should be applicable across all branches of surgery, particularly if both generic and condition-specific adverse events were selected. This would then allow, or be more likely to identify in the same articles, a direct link to the monitoring of adverse events rather than of operations and procedures.

There are also different approaches to the monitoring of adverse events. At one extreme is the critical incident approach, by which individual surgeons agree to identify single events that are then reviewed and discussed. This may be done locally, as in morbidity and mortality meetings, or nationally, as in confidential enquiries into deaths. At the other extreme is the identification of events from data collection or record systems, which may or may not necessarily be collected for that purpose, usually with a view to calculating incidence rates and trends over time. Examples include hospital nosocomial programmes, the Health Care Finance Administration (HCFA) publication of death rates after coronary artery bypass graft (CABG), or some of the Scottish Clinical Outcomes Indicators. In between are various combinations, such as intermittent or continuous surgical audits on a regional or national scale. It was important to attempt to review the range of approaches because the attributes of individual systems would vary and would be likely to involve different levels of validity and reliability. A particular issue would be the ability to monitor time trends and compare rates, and therefore the volume of expected events was also important. Moreover, it was possible that there would be more literature about common events than about events, other than death, that were less common but not so rare as to cause special interest. It was therefore decided to examine frequency of likely occurrence before making a selection of topics. This was done by looking at the four UK national routine data sets. Also, on the assumption that existing surgical audits might yield information on both the nature and the frequency of surgical adverse events that might not be found in the published literature, the
coordinators of these audits were approached. Thus, once the general approach to the review had been decided upon, the next stage was to select surgical adverse events for inclusion.

**National surgical data**

Hospital-based data were purchased from the four national sources to identify the most frequently occurring postoperative complications at time of discharge from hospital and at time of death in hospital (Box 1). Data on the most commonly performed surgical procedures by volume and their surgical complications were also requested. The data obtained from the four national sources were not directly comparable as they were in various formats, with different International Classification of Diseases (ICD) codes used over different time periods. However, it was sufficient to give the review group an indication of the most frequently performed operative procedures throughout the UK and related complications at time of hospital discharge (Box 2). The most frequently recorded postoperative complications identified from national sources included: mechanical complications of internal devices, implants and grafts; postoperative infection; haemorrhage and haematoma complicating a procedure; phlebitis and thrombophlebitis of the lower extremities; and cardiac and respiratory complications.

Written requests also were sent to the Royal Colleges of Surgery and a number of surgical societies and associations for information on large-scale, regional or national surgical audits that included surgical complications or adverse events as outcome measures (see Box 1). Information was obtained in the form of national audit reports, scientific publications, surgical audit data collection forms and unpublished data from ongoing national surgical audits. Surgical audits at department or ward level were excluded. The surgical complications that were recorded and monitored varied according to surgical specialty, but were broadly similar to those recorded and reported in the national statistics. Examples included wound, chest and urine infections, complications of the deep vessels of lower extremities, haemorrhage (intraoperatively), pulmonary embolism, anastomotic leak, cardiac complications and renal failure.

**Rationale for rejection of certain surgical adverse events**

The review team selected examples of postoperative events based on the frequency of occurrence from the national hospital discharge data, the surgical audit literature and after review of surgical mortality annual audit reports. This process involved lengthy discussion and deliberation during the early stages of the study, during which a number of ‘common’ postoperative events were excluded. Fundamentally, given the focus on linking to monitoring, it was not clear that they would have added to the generalisable lessons to be learned, and to include them would have meant reducing the depth of coverage of the other items. For example, a decision was made to exclude cardiac complications (e.g. myocardial infarction) because of the potential difficulty in attributing these adverse events to the surgical intervention, given their overall frequency in the general population. Postoperative chest, urine and wound infections were prevalent events and were also known to be recorded within hospital nosocomial monitoring systems. We were less confident that chest infection and urine infection were directly attributable to the antecedent surgical procedure, unlike surgical wound infection.

Although ‘haemorrhage or haematoma complicating a procedure’ was commonly recorded at time of discharge, the ICD-10 code did not distinguish timing of occurrence. Haemorrhage can occur perioperatively (primary haemorrhage), during the immediate postoperative period (early) or have delayed presentation (secondary

---

**BOX 1 National data**

**Sources of national data**


**Surgical associations and societies**

British Association of Otorhinolaryngologists Head & Neck Surgeons

British Association of Oral & Maxillofacial Surgeons

British Association of Paediatric Surgeons

British Association of Plastic Surgeons

British Association of Urological Surgeons

British Orthopaedic Association

The Society of Neurological Surgeons

The Society of Cardiothoracic Surgeons of Great Britain & Ireland

The Vascular Surgical Society
Haemorrhage occurring during the immediate postoperative period usually indicates inadequate operative haemostasis or a technical mishap, such as slipped ligature or unrecognised trauma to a blood vessel, whereas later or delayed haemorrhage occurs several days post-operatively and is often related to infection that has eroded vessels at the operative site. In 1997, the Scottish Audit of Surgical Mortality (SASM) highlighted the difficulty in attributing death to haemorrhage, even when investigated at the level of the individual. Thus haemorrhage as an adverse event was excluded from the review. Following analysis, interpretation and discussion of routine and audit data, the review group selected four surgical adverse events on the basis of their frequency and burden of illness, and to represent a spectrum of information on definition, measurement and monitoring.

**BOX 2** Top ten complications (not ranked) of the ICD-9 and ICD-10 codes, and the procedures and most commonly occurring associated complications of OPCS-4

<table>
<thead>
<tr>
<th>ICD-9 code</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>998.5</td>
<td>Postoperative infection</td>
</tr>
<tr>
<td>998.1</td>
<td>Haemorrhage or haematoma complicating a procedure</td>
</tr>
<tr>
<td>996.6</td>
<td>Mechanical complication of internal orthopaedic device, implant or graft</td>
</tr>
<tr>
<td>996.6</td>
<td>Infection and inflammatory response to internal prosthetic device, implant or graft</td>
</tr>
<tr>
<td>998.3</td>
<td>Disruption of operation wound</td>
</tr>
<tr>
<td>451.1</td>
<td>Of deep vessels of lower extremities</td>
</tr>
<tr>
<td>997.1</td>
<td>Cardiac complications</td>
</tr>
<tr>
<td>997.3</td>
<td>Respiratory complications</td>
</tr>
<tr>
<td>998.8</td>
<td>Other specified complications not elsewhere classified</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>ICD-10 code</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>T81.4</td>
<td>Infection following a procedure not elsewhere classified</td>
</tr>
<tr>
<td>T81.0</td>
<td>Haemorrhage or haematoma complicating a procedure not elsewhere classified</td>
</tr>
<tr>
<td>T84.0</td>
<td>Mechanical complication of internal joint prosthesis</td>
</tr>
<tr>
<td>I80.2</td>
<td>Phlebitis or thrombophlebitis of other deep vessels of lower extremities</td>
</tr>
<tr>
<td>T81.3</td>
<td>Disruption of wound not elsewhere classified</td>
</tr>
<tr>
<td>T85.2</td>
<td>Mechanical complication of intraocular lens</td>
</tr>
<tr>
<td>I97.8</td>
<td>Other postprocedural disorders of circulatory system not elsewhere classified</td>
</tr>
<tr>
<td>T82.2</td>
<td>Mechanical complication of coronary artery bypass or valve grafts</td>
</tr>
<tr>
<td>J95.8</td>
<td>Other postprocedural respiratory disorders</td>
</tr>
<tr>
<td>N99.8</td>
<td>Other postprocedural disorders of genitourinary system not elsewhere classified</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>OPCS-4*</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>W37</td>
<td>Total prosthetic replacement of hip joint using cement: mechanical complication of internal orthopaedic device, implant or graft</td>
</tr>
<tr>
<td>K40</td>
<td>Saphenous vein graft replacement of coronary artery: cardiac complications</td>
</tr>
<tr>
<td>X55</td>
<td>Other operations on unspecified organ: of deep vessels of lower extremities</td>
</tr>
<tr>
<td>S57</td>
<td>Exploration of other skin of other site: postoperative infection</td>
</tr>
<tr>
<td>C75</td>
<td>Prosthesis of lens: mechanical complication of other specified prosthetic device, implant or graft</td>
</tr>
<tr>
<td>Q07</td>
<td>Abdominal excision of uterus: haemorrhage or haematoma complicating a procedure</td>
</tr>
<tr>
<td>S47</td>
<td>Opening of skin: postoperative infection</td>
</tr>
<tr>
<td>T30</td>
<td>Opening of abdomen: haemorrhage or haematoma complicating a procedure</td>
</tr>
<tr>
<td>L91</td>
<td>Other vein related operations: infection and inflammatory reaction due to internal prosthetic device, implant or graft</td>
</tr>
<tr>
<td>F34</td>
<td>Excision of tonsil: haemorrhage or haematoma during a procedure</td>
</tr>
<tr>
<td>H33</td>
<td>Excision of rectum: gastrointestinal complications</td>
</tr>
<tr>
<td>F10</td>
<td>Simple extraction of tooth: haemorrhage or haematoma complicating a procedure</td>
</tr>
</tbody>
</table>

*Procedure, followed by the most commonly occurring associated complication
Rationale for inclusion of surgical wound infection
Surgical wound infection is a common, generic, postoperative event that causes considerable morbidity but seldom leads to death. Surgical wound infection can potentially occur after any surgical incision, regardless of whether or not minimally invasive techniques are employed. Up to 15% of all elective surgical patients will develop a surgical wound infection, although rates as high as 30% are not uncommon after contaminated or dirty surgical procedures. The implications of postoperative wound infection include patient morbidity and suffering, risk to other patients, additional use of resources, and delayed discharge in the presence of severe infection. Surgical wound infection was recorded in large-scale surgical audits and frequently occurred across the four national routine data sets (ICD-10 codes T81.4 and T81.3). Given the frequency of the event, it was anticipated that a large body of published literature related to definition and measurement would be found. Surgical wound infection was also known to be recorded within nosocomial monitoring programmes, both in the UK and abroad. Although other infections, such as postoperative chest and urine infection, were also prevalent events, the review group was less confident that these events could be directly attributable to the antecedent surgical procedure, unlike wound infection. Useful lessons could also be learned from the recent establishment of the English and Welsh hospital infection monitoring systems. Finally, wound infection was highlighted as an event in the original brief by the HTA Programme.

Rationale for inclusion of anastomotic leak
Anastomotic leak is the breakdown of the operative union of two hollow or tubular structures. Anastomotic breakdown is an important complication after gastrointestinal surgery and can lead to severe and fatal consequences. It is associated with increased morbidity and mortality and prolonged hospital length of stay, and can also impact upon long-term outcome.15 Anastomotic leak is an uncommon event and was not ranked among the complications at time of hospital discharge from the national sources, although it is possible that the generic complication ‘gastrointestinal complication’ after ‘excision of rectum’ (Office of Population Censuses and Surveys (OPCS) 4 code H33) might include leak. Anastomotic leak was selected as it had been consistently identified by the SASM as a cause of significant mortality, with the risk of fatality from anastomotic leak associated with a delay in recognition, detection and investigation.14 Anastomotic leak, therefore, was regarded by the review group as an important and potentially modifiable and avoidable surgical adverse event. Furthermore, this event is specific to gastrointestinal surgery rather than generic, and is relatively uncommon compared to surgical wound infection and deep vein thrombosis (DVT).

Rationale for inclusion of DVT
DVT is a common, generic postoperative event. Early diagnosis of postoperative DVT is essential to reduce the risk of pulmonary embolism and incidence of post-thrombotic sequelae (e.g. chronic pain, recurrent cellulitis, deep vein insufficiency). The incidence of DVT following orthopaedic surgery is particularly high, with a 20% risk of developing DVT after hip fracture surgery, even when effective prophylactic regimens are implemented.15 DVT was one of the top ranked postoperative events recorded in the UK national statistics (ICD-10 code I80.2) and also was recorded in a number of national and/or large-scale surgical audits. The importance of a valid definition and accurate methods for diagnosis and measurement of DVT was highlighted in the original brief by the HTA Programme.

Rationale for inclusion of surgical mortality
Surgical mortality is an uncommon event but can occur after any surgical procedure. Approximately 3500 patients in Scotland die each year within 30 days of surgery, the equivalent number in England and Wales being 20,000 patients.16,17 Surgical mortality was chosen because of the effort that was known to have gone into developing comprehensive systems for monitoring surgical mortality, both within the UK and internationally. The national UK mortality monitoring systems, the National Confidential Enquiry into Peri-Operative Deaths (NCEPOD) and the SASM, both aim to identify remedial factors, namely the factors of care that might have delayed or prevented death.16,17 The review panel was also aware that surgical death was recorded by cardiac and vascular registries. By including surgical mortality as an event it would be possible to concentrate on fundamental issues of ascertainment and attribution, such as timing of follow-up, rather than on the definition ‘death’ itself.
The four selected surgical adverse events were, therefore, surgical wound infection, anastomotic leak, DVT and mortality. These events covered a range in terms of burden of illness (low morbidity, high morbidity, fatal), epidemiology (frequent, infrequent) and specificity (generic, surgery specific) (Table 1).

**R**emit and coverage of the review

The remit of the study was to review the definition, measurement and monitoring of surgical adverse events. This was not an attempt to identify all published articles related to selected adverse events regardless of quality, but rather a systematic and comprehensive overview of methodology related to the definition and/or measurement and/or monitoring for each event. Table 2 highlights the literature covered for each included surgical adverse event. The group had to decide what was to be extracted from individual studies. For each event, we were interested in the quality of the recording of each event (e.g. if a definition was used, the way in which it was measured) as well as the quality of each study, although the two are related.

**Assessment criteria for definition and measurement**

Explicit criteria for evaluation and critical appraisal of the evidence on definition, measurement and monitoring were determined at an early stage in the review. For definition and measurement the following properties were assessed:

- validity (content, criterion, construct)
- reliability
- acceptability and practicality.

**Validity**

Validity is an assessment of the extent to which something measures what it purports to measure.18 An observation is valid if it corresponds to the true state of the phenomenon being measured. It is important to remember that a definition or measurement of an adverse event may be valid only with respect to the patient group it was evaluated on and cannot necessarily be extrapolated to other categories. A number of components of validity require to be satisfied: content (face) validity, criterion validity and construct validity.

**Content validity**

Any assessment of a surgical adverse event must have face validity; that is, it must appear to assess the relevant properties of the event. We examined the content of the definition for clinical sense and the method of measurement for relevance to the nature of the adverse event. For example, a grading scale would be invalid for an event with no range of severity.

**Criterion validity**

Criterion validity has been defined as the correlation of an event or scale with some other measure of the event, ideally the gold standard, which has been used and accepted in the field.19

---

**TABLE 1** The features of surgical adverse events included in the review

<table>
<thead>
<tr>
<th>Attributes</th>
<th>Surgical wound infection</th>
<th>Anastomotic leak</th>
<th>DVT</th>
<th>Mortality</th>
</tr>
</thead>
<tbody>
<tr>
<td>Site of surgery</td>
<td>Generic</td>
<td>Gastrointestinal</td>
<td>Generic</td>
<td>Generic</td>
</tr>
<tr>
<td>Burden of illness</td>
<td>Low</td>
<td>High</td>
<td>Medium/high</td>
<td>Fatal</td>
</tr>
<tr>
<td>Epidemiology</td>
<td>Frequent</td>
<td>Rare</td>
<td>Frequent</td>
<td>Rare</td>
</tr>
<tr>
<td>Risk of fatality</td>
<td>Low</td>
<td>High</td>
<td>Medium/high</td>
<td>Variable</td>
</tr>
</tbody>
</table>

**TABLE 2** Coverage of literature for selected surgical adverse events

<table>
<thead>
<tr>
<th>Surgical adverse event</th>
<th>Definition</th>
<th>Measurement</th>
<th>Surveillance/monitoring systems</th>
</tr>
</thead>
<tbody>
<tr>
<td>Surgical wound infection</td>
<td>✔</td>
<td>✔</td>
<td>✗</td>
</tr>
<tr>
<td>Anastomotic leak</td>
<td>✔</td>
<td>✔</td>
<td>✗</td>
</tr>
<tr>
<td>Surgical mortality</td>
<td>✔</td>
<td></td>
<td>✗</td>
</tr>
<tr>
<td>DVT</td>
<td>✔</td>
<td></td>
<td>✗</td>
</tr>
</tbody>
</table>
Methods

For assessment of surgical adverse events, criterion validity was assessed by comparing different definitions or measurement techniques applied to the same population at the same time (e.g. a comparison of scoring systems for surgical wound infection, a comparison of diagnostic tests for DVT). Details on sensitivity and specificity were extracted from each paper or calculated by the review group if the data were published. The methods used to calculate values when assessing diagnostic utility are given in full in appendix 2.

Construct validity
Construct validity is examined by quantitatively examining the relationships of a construct to a set of other variables. Fitzpatrick and co-workers state that “no single observation can prove the construct validity of a new measure; rather it is necessary to build up a picture from a broad pattern of relationships of the new measure with other variables”. Assessment is based on whether a construct behaves in a way you would expect it to if it truly reflected the event under study. In terms of the definition and measurement of surgical adverse events, construct validity would be judged by whether, for example, a definition of wound infection reflected an expected clinical response to an antibiotic.

Reliability
Reliability is an assessment of the extent to which the definition and measurement of the surgical adverse event is repeatable and reproducible. Both objective and subjective criteria for the identification of a surgical adverse event are subject to random error. If a method has poor repeatability, this will lead to poor agreement between observers or different methods of measurement. Information was extracted on whether the measurement was repeated on more than one occasion and/or by more than one observer and estimates of repeatability (intra-rater reliability) and reproducibility (inter-rater reliability) sought. The extent of agreement, if reported, was extracted from individual studies. The statistical methods for assessment of inter- and intra-rater reliability are given in full in appendix 2.

Acceptability and practicality
A definition and measurement must be acceptable, in that it is easy to comprehend and suitable for use in clinical practice. Any comments on the practicality of using the definition or the method of measurement were extracted, including comments on use of time, resources and personnel.

Assessment criteria for monitoring systems

The term ‘monitor’ has been defined as “persons or devices for checking or warning about a situation” and “to maintain regular surveillance over”. The monitoring of a surgical adverse event was interpreted by the review group as any information or surveillance system where data were collected on the selected surgical adverse event in a repeated and systematic manner. Surveillance has been described as the “the on-going systematic collection, analysis, evaluation and dissemination of data”. We were interested in particular epidemiological features of relevant information and surveillance systems:

- The coverage of the system: whether the information system was national, regional, hospital or unit based; the proportion of ‘units’ included; and estimated coverage within each unit. Variation in levels of completion by individual participants within units were also included wherever possible.
- Whether or not denominator data were collected of the population at risk to allow calculation of incidence rates of the selected adverse event.
- Whether the information system included a definition of the selected adverse event and whether it was a standard and agreed definition. The same criteria for evaluation of definition and measurement were applied in the assessment of monitoring systems.
- Whether the information system allowed for assessment of risk factors for individuals or groups of patients. If so, what risk factors were included (e.g. operation time, event classification, American Society of Anesthesiology (ASA) score, co-morbidity).
- Issues related to data collection, in terms of who collected data, what methods were used, how data were collated and whether validation checks were conducted.
- When the event was first recorded and entered into the monitoring system. At what stage in the ‘natural history’ of the event it was recorded on the monitoring system and details of follow-up postdischarge (postdischarge surveillance and monitoring).
- Output from the system, both in terms of feedback to individual surgeons or groups of surgeons, and wider dissemination to the academic and clinical field. Information was extracted on the impact and costs of the systems, where available.
Literature search

Bibliographic databases

Electronic search strategies were developed by one researcher (JB) and are listed in appendix 1. Searching for published literature was undertaken on the main health and biomedical bibliographic databases: MEDLINE, EMBASE, HealthSTAR, CINAHL and the Cochrane Library. Further searches were also made on the internet and PubMed at a later stage in the review. Preliminary searches were made of the Science Citation Index and Edinburgh Data & Information Access (EDINA) databases, but were discontinued at an early stage in the review because of the low yield of relevant material. It was not possible to identify a small number of ‘key’ journals to handsearch as our selected adverse events were not specific to any operative procedure or specialty. Handsearching of key British surgical journals (British Journal of Surgery, Annals of the Royal College of Surgeons of England) was conducted for a four-year period for one adverse event (surgical mortality information systems) to estimate the usefulness of this process, but was discontinued due to a low return of relevant literature.

Other sources

Contact was made with surgical colleges, societies and associations at an early stage in the review for details of unpublished studies related to the measurement and monitoring of surgical complications. Requests for unpublished literature relating to the monitoring of surgical wound infection were made to senior infection personnel at the Public Health Laboratory Service (PHLS) and the Scottish Centre for Infection and Environmental Health. Although various reports were obtained, these related to hospital rates of wound infection rather than wound infection. A decision was made to widen the search to include all prospective, cohort, follow-up and longitudinal studies with ‘surgical wound infection’ as an MeSH term and to extract details on definition and measurement from literature published over a specified time period. Fifteen separate search strategies were developed for the topic of surgical wound infection alone. Initial searches for literature on the validity and reliability of definition and measurement were conducted for the period 1985–1999, but less than 20 articles were found. The review team made the decision to extract details on definition and measurement from literature published over a specified time period. Given the large volume of prospective studies on surgical complications and multiple search strategies (n = 30), searching was restricted to studies published in the English language only over a 7-year period, from 1993 to 1999. Because of the large volume of English-language literature, translation of foreign language articles was not undertaken, although many examples of international monitoring systems were identified (published in English). Inclusion and exclusion criteria for each separate adverse event is expanded on below.

Study design

In searching for evidence of reliable and valid definitions and their measurement, the search was limited to prospective studies as it was more likely that the adverse event would be defined before data collection and retrospective studies do not permit the accurate identification of denominator data. However, when the study design and aim were not clear from the abstract, the article was photocopied or retrieved for review. Although the search for definitions and measurement was limited to prospective studies, if a study was described as a review and it was clear that agreed definitions had been used throughout then it was included. Similarly, if prospective data collection had been conducted using standard definitions and methods, but data were analysed at a later stage and presented as a retrospective review, the study was included in the review. The restriction on prospective literature was lifted when searching for epidemiological features of surgical monitoring systems, as reports were often descriptive or narrative and application of prospective terms (e.g. cohort, follow-up study) severely limited the quality and scope of available literature.

Language and time limits

Initial searches for literature on the validity and reliability of definition and measurement were conducted for the period 1985–1999, but less than 20 articles were found. The review team made the decision to extract details on definition and measurement from literature published over a specified time period. Given the large volume of prospective studies on surgical complications and multiple search strategies (n = 30), searching was restricted to studies published in the English language only over a 7-year period, from 1993 to 1999. Because of the large volume of English-language literature, translation of foreign language articles was not undertaken, although many examples of international monitoring systems were identified (published in English). Inclusion and exclusion criteria for each separate adverse event is expanded on below.

Inclusion and exclusion criteria

Surgical wound infection

Fifteen separate search strategies were developed for the topic of surgical wound infection alone. Initial searches had a good return on surveillance and monitoring literature but very few articles were found on the definition and measurement of wound infection. A decision was made to widen the search to include all prospective, cohort, follow-up and longitudinal studies with ‘surgical wound infection’ as an MeSH term and to extract details on definition and measurement from individual studies. Three search strategies were designed to retrieve prospective, measurement- and monitoring-related literature across five databases. The term ‘wound infection’ was narrowed to ‘surgical wound infection’ in an
attempt to exclude wounds caused by gunshot, trauma, burns, pressure sores and non-surgical injury. Many surgical procedures do not result in conventional surgical wounds (e.g. vaginal hysterectomy, transurethral prostatectomy, myringoplasty) or wounds involving a skin incision (e.g. replacement of intraocular lens). The panel therefore excluded a number of articles from different surgical specialties (ophthalmology, ear nose and throat surgery, urology, obstetrics and gynaecology) from the wound infection section. Articles that included an assessment of wounds from a range of surgical procedures, such as general and gynaecology surgery, were included for review. A later decision was made to exclude orthopaedic surgery, because infection was often related to rejection of a prosthesis or implant and had different signs, symptoms and presentation from those of surgical wound infection, which occurred in the early and intermediate postoperative period. The subsequent consideration of surgical wound infection, however, is relevant to orthopaedic surgery. Standard or widely accepted definitions of surgical wound infection and anastomotic leak in the surgical literature published before 1993 were traced and included in the review. Each of these articles published before 1993 has been clearly documented and described in chapters 3 to 5.

**Anastomotic leak**

Although MEDLINE included an MeSH term for surgical anastomosis (Anastomosis, Surgical), unlike EMBASE, there was no specific term for anastomotic leakage, breakdown or dehiscence. Following group discussion, a decision was made to expand the literature search to include and review all abstracts with ‘Anastomosis, Surgical’ as an MeSH term. The search was modified by exchanging measurement terms with those related to study design (prospective, longitudinal, follow-up and cohort studies) to generate a larger body of literature. This modification resulted in over 1900 abstracts for review. Studies of gastrointestinal or hepatobiliary anastomoses (e.g. Roux-en-Y anastomosis, cholecystostomy, choledochoestomy, gastroenterostomy, jenunoileal bypass, hepatic portoenterostomy, pancreatojejunostomy) were eligible for inclusion in the review. A large number of abstracts related to non-gastrointestinal anastomoses, such as vascular and neurosurgical shunts and implants (e.g. arteriovenous shunts, cerebrospinal fluid shunts, endolymphatic shunts, peritonovenous shunts, portasytesticshunts, heart bypass shunts, salpingoscopies, vasovasotomies). These were excluded from the review.

**Surgical mortality**

Preliminary searching for articles related to surgical mortality was undertaken as an entry to identifying surgical monitoring systems. This section proved challenging as the search generated a large number of heterogeneous articles, including audits, research studies, descriptions of surgical monitoring systems and mortality-related risk adjustment and risk scoring systems. The strategy comprised subject headings and text words relating to information systems, databases and monitoring terms combined with surgical mortality. The strategy was modified to include prospective study design terms (e.g. prospective, cohort, follow-up, longitudinal), but these were found to be too restrictive. They were, therefore, removed and the original searches used. Although the majority of abstracts reviewed were prospective in nature, a number were retrospective analyses or descriptions of prospectively collected data from registries or systems and were obtained in print for review. The review panel was aware that this decision to include narrative and non-prospective studies differed from the three other adverse events but, given that the focus was on the monitoring of surgical mortality and examples of systems, restriction to prospective studies would severely limit the quality and scope of available literature. The types of articles deemed eligible included:

- examples of surgical monitoring systems, other than national routine data collection systems, with a primary aim to identify, record and monitor surgery-related mortality
- examples of surgical monitoring systems, either procedure specific, condition specific or related to surgical specialty, that recorded mortality as an outcome but as a ‘by-product’ rather than as the main purpose of the system.

Numerous surgical audits have been published but, for the purposes of this review, single reports, ward-, department- or unit-based audits were excluded unless they were examples of sustained, ongoing projects. Articles specific to risk assessment and risk scoring were excluded from the review. Although acknowledged as integral for intra- and inter-institutional comparisons of mortality rates, they would require a separate review strategy. Chapters 8 and 9, which discuss surgical mortality monitoring systems, include examples of ‘major’ reports that were published before 1993 because they contributed to the overall understanding of the monitoring of mortality and the system in question.
Deep vein thrombosis
Searches for literature on DVT were undertaken to identify studies of the properties of measurement, including assessment of the validity, reliability, accuracy, sensitivity and specificity of diagnostic tests. Randomised controlled trials of surgical interventions or therapies that reported postoperative incidence or prevalence of DVT were excluded unless the citation contained MeSH terms of measurement. This focused the search on studies of imaging techniques and comparisons between different diagnostic tests and other methods of detecting DVT. The search strategy was expanded to include all literature related to the diagnostic accuracy of the measurement of DVT rather than restricted solely to the surgical or postoperative literature.

Abstract review and management of references
All abstracts were printed and read independently by two assessors. When reviewers were unsure of whether or not an abstract was eligible for inclusion, a hard copy of the article was obtained.

Table 3 details the total volume of abstracts printed and read from each electronic bibliographic database. It is possible within HealthSTAR to limit abstracts to those not cited on MEDLINE, but this facility is not available for EMBASE, CINAHL or the Cochrane Library. A number of citations, therefore, were duplicated as they appeared on more than one database.

All references retrieved from electronic sources were downloaded and stored in the bibliographic management software Reference Manager Professional Network Edition (version 9.0). Whenever possible, the source of the reference was added to each citation. References were initially stored in separate databases according to subject.

Data extraction forms
Two data collection forms were designed for reviewers to record extracted data (appendix 2). Prospective forms comprised sections to record the following details: bibliographic details, unique identifier, reviewer, study design, surgical specialty, surgical procedure/intervention, search details, adverse event of interest, definition of event, grading or scoring system, references to definition or system, clinical factors considered in assessment, and investigations used to determine an event. A separate section was included for articles of monitoring and surveillance systems for details of scale, coverage, definition, data collection, entry, analysis, presentation and user feedback. Measurement forms comprised sections on definition, validity, accuracy, sensitivity, specificity, inter- and intra-rater reliability, practicality and feasibility of the measure or diagnostic test. Two reviewers independently assessed each study and discrepancies were by discussion between reviewers or by discussion and further review with other panel members.

TABLE 3 Abstracts obtained per database for surgical adverse events

<table>
<thead>
<tr>
<th>Surgical adverse event</th>
<th>MEDLINE*</th>
<th>EMBASE†</th>
<th>CINAHL*</th>
<th>HealthSTAR*</th>
<th>Cochrane Library‡</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>Surgical wound infection</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>All</td>
<td>2010</td>
<td>1020</td>
<td>328</td>
<td>32</td>
<td>393</td>
<td>3783</td>
</tr>
<tr>
<td>Anastomotic leak</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>All articles</td>
<td>817</td>
<td>855</td>
<td>154</td>
<td>2</td>
<td>82</td>
<td>1910</td>
</tr>
<tr>
<td>DVT</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Measurement</td>
<td>1109</td>
<td>873</td>
<td>31</td>
<td>7</td>
<td>522</td>
<td>2542</td>
</tr>
<tr>
<td>Surgical mortality</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Information systems</td>
<td>1285</td>
<td>1985</td>
<td>204</td>
<td>28</td>
<td>253</td>
<td>3755</td>
</tr>
</tbody>
</table>

* Retrieval software OVID Technologies
† Retrieval from BIDS via the internet using the OVID interface
‡ The Cochrane Library, CD-ROM, Issues 2 and 3
Chapter 3

The definition and measurement of surgical wound infection

Surgical wound infection, although it seldom causes mortality, leads to morbidity and prolonged hospital stay. The process of development of postoperative wound infection is known to be multifactorial in nature and arises from the complex interaction of host and environment factors. In 1976, Altemeier and co-workers defined the classic signs of infection as "the invasion and multiplication of micro-organisms in body tissues that result in local cellular injury" and described it as "the inclusion of redness, tenderness, warmth, swelling of the surrounding skin and presence of pus". This chapter summarises the various definitions of surgical wound infection and the different methods used to measure wound infection.

Definition

A total of 112 prospective studies were critically appraised and details were extracted on the definition and measurement of surgical wound infection. Thirty-one studies were excluded from the review for the following reasons:

- no definition of surgical wound infection included in study (n = 23)
- retrospective study design (n = 4)
- wound infection not reported in study (n = 2)
- meta-analyses without definitions from primary studies (n = 2).

Of the 31 excluded studies, 26 reported surgical wound infection rates but did not define wound infection. Many of these studies stated that the primary aim was to measure the incidence of postoperative wound infection, yet descriptions of what was measured and how it was measured were vague. For example, one randomised controlled trial compared groups with regard to the incidence of "wound infection, postoperative temperature and length of hospital stay", but no description of the definition or data collection was given. Similarly, a trial of laparoscopic versus open appendectomy was conducted to "determine the incidence and nature of post-operative infectious complications" and, although wound infection rates are presented, the definition of wound infection was not described in the article. Two meta-analyses were excluded as they did not include a description of the definitions or measurement of wound infection from the primary studies. The reference lists from both publications were used to trace and retrieve primary studies not identified from the original search.

A total of 82 studies fulfilled the appraisal criteria in that they contained a definition and described the measurement of surgical wound infection. These studies used a variety of different definitions, which ranged from simple descriptions, such as ‘the presence of pus’ to criteria that distinguished levels and components of surgical wound infection.

Five nationally proposed definitions were identified from the literature search. Three were published by groups within the UK: Glenister and co-workers from the PHLS; the Surgical Infections Society Group (SISG); and the Second UK National Prevalence Survey (NPS). Two definitions were identified from the Centers for Disease Control (CDC) in the USA. These five definitions were each published as the result of multidisciplinary or consensus groups working in the field of surgical wound infection. A full description and comparison of the five standardised definitions is given before expanding on the other definitions identified from the literature review.

Glenister and co-workers

This definition, published by the PHLS in 1992, was used in a UK study of the effectiveness of different surveillance methods in detecting hospital infections (Box 3). It was designed for use by infection control nurses.

The Surgical Infection Study Group

In 1991, the SISG expressed a need to establish definitions of infectious morbidity for clinical application by surgeons from different surgical specialties. The definition of surgical wound infection given in Box 4 was one of a number of
The definition and measurement of surgical wound infection

separate postoperative infection definitions proposed during a 1990 workshop.

**The Second UK NPS**

The Steering Group of the Second UK NPS of hospital acquired infections published definitions in 1993. The Joint Working Party Steering Group who agreed on the definitions comprised personnel from the Hospital Infection Society, the Laboratory of Hospital Infection Colindale and the Infection Control Nurses’ Association. These definitions are identical, with the exception of timing, to those published by the SISG in 1991, but are presented as itemised points (Box 5).

**The Centers for Disease Control**

Two definitions of surgical wound infection were developed by the CDC in Atlanta, first in 1988 (Box 6) and a later modified version in 1992 (Box 7). The definition for surgical wound

**BOX 3 Definition of surgical wound infection: Glenister and co-workers***

A *wound* is defined as a break in the epithelial surface (skin or mucous membrane) and the underlying tissue made by some positive act, such as an accident or surgical incision. Burns should be excluded. An ulcer or pressure sore is not a wound for the purposes of this definition. All *wound infections* must have purulent discharge in or exuding from a wound or seen on direct examination at the operative site.

**Major infection** is present when the wound is broken down, gaping or completely dehisced, or there is evidence of septicemia, spreading cellulitis or lymphangitis.

**Minor infection** is present when the wound is not broken down, gaping or completely dehisced and there is no evidence of septicemia, spreading cellulitis or lymphangitis.

**Surgical wound infection**

Infection occurs at the incision or operative site (including drains) within 30 days after surgical operation if no implant is left in place or within one year if an implant is left in place. The infection must appear to be related to the surgical procedure.

**Accidental wound infection**

Infection occurs at or in the accidental wound site.

**Note:** Infections occurring at the entry site of a device that has required an incision for insertion should be noted as surgical wound infection (e.g. tracheostomy, intravascular catheters, renal dialysis catheters, suprapubic catheter). The presence of the device should be noted.

**BOX 4 Definition of surgical wound infection: the SISG***

**Wound infection**

A *wound* is defined as a break in an epithelial surface, which may be surgical or accidental. Burns, ulceration and pressure sores have been excluded in this definition, but drain sites should be included. A wound infection should have either a purulent discharge in, or exuding from, the wound, or a painful, spreading erythema indicative of cellulitis.

Bruising, haematoma formation, and serous and lymph collections are complications that may predispose to the development of wound infection, and may lead to diagnostic difficulties. Infection should be considered to be present when there is fever, tenderness, oedema and an extending margin of the erythema. The discharge of clear fluid from a wound does not indicate an infection unless it is accompanied by cellulitis. The definition of *wound infection* should not be dependent on the results of bacteriological studies. False-negative cultures can occur, and on other occasions organisms isolated from cultures may represent either secondary colonisation or merely contamination. Wound infection may be classified according to aetiology, time or severity.

**Primary and secondary wound infection**

The infection should be considered *primary* unless there is a predisposing complication. *Secondary* infection may follow a complication that results in the discharge of serum, haematoma, cerebrospinal fluid, urine, bile, pancreatic juice, gastric or intestinal contents from the wound, contaminated by bacteria from within the patient or from the environment.

**Time**

With regard to time, wound infection may be divided into:

- **early**, presenting within 30 days of operation
- **intermediate**, presenting between 1 and 3 months after operation
- **late**, presenting more than 3 months after operation.

**Severity**

Wound infection should be classified as:

- **minor**, when there is discharge of pus from the wound without lymphangitis or deep tissue destruction
- **major**, when the purulent discharge is accompanied by partial or complete dehiscence of the fascial layers of the wound or by spreading cellulitis and lymphangitis that requires antibiotic therapy.
infection was part of a set of definitions used for the surveillance of nosocomial infection.\textsuperscript{31} The new definitions were introduced to hospitals participating in the National Nosocomial Infections Surveillance (NNIS) system. In 1992, the definitions were reviewed and modified by the Surgical Wound Infection Task Force.\textsuperscript{32} This task force comprised members from the Society for Hospital Epidemiology of America, the Association for Practitioners in Infection Control, the Surgical Infection Society and the CDC.\textsuperscript{32} The term ‘surgical wound infection’ was changed to ‘surgical site infection’, in line with surgical terminology, and a clearer distinction made between superficial and deep infections of the incision. The 1992 modification was more

<table>
<thead>
<tr>
<th>BOX 5 Definition of surgical wound infection: Second UK NPS\textsuperscript{30}</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Wound infection</strong></td>
</tr>
<tr>
<td>A wound is defined as a break in an epithelial surface, which may be surgical or accidental. A wound infection should have either a purulent discharge in or exuding from the wound, or a painful, spreading erythema indicative of cellulitis. Infection should be considered to be present when there is fever (&gt; 38°C), tenderness, oedema and an extending margin of erythema, or the patient is still receiving active treatment for a wound that has discharged pus.</td>
</tr>
<tr>
<td>• Bruising, haematoma formation, and serous and lymph collections are complications that may predispose to the development of wound infection, and may lead to diagnostic difficulties.</td>
</tr>
<tr>
<td>• The discharge of clear fluid from a wound does not indicate an infection unless it is accompanied by cellulitis.</td>
</tr>
<tr>
<td>• Wound infection may be classified according to aetiology, time or severity.</td>
</tr>
<tr>
<td>• Wound infection should be classified as <strong>minor</strong> when there is discharge of pus from the wound without lymphangitis or deep tissue destruction, and <strong>major</strong> when the purulent discharge is accompanied by partial or complete dehiscence of the fascial layers of the wound or by spreading cellulitis and lymphangitis that requires antibiotic therapy.</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>BOX 6 Definition of surgical wound infection: CDC, 1988\textsuperscript{31}</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Incisional surgical wound infection</strong> must meet the following criteria: infection occurs at incision site within 30 days after surgery and involves skin, subcutaneous tissue, or muscle located above the fascial layer and any of the following:</td>
</tr>
<tr>
<td>• Purulent drainage from incision or drain located above fascial layer.</td>
</tr>
<tr>
<td>• Surgeon deliberately opens wound, unless wound is culture-negative.</td>
</tr>
<tr>
<td><strong>Deep surgical wound infection</strong> must meet the following criteria: infection occurs at operative site within 30 days of surgery if no implant* is left in place or within one year if an implant is in place and infection appears related to surgery and infection involves tissues or spaces at or beneath the fascial layer and any of the following:</td>
</tr>
<tr>
<td>• Purulent drainage from drain placed beneath fascial layer.</td>
</tr>
<tr>
<td>• An abscess or other evidence of infection seen on direct examination, during surgery, or by histopathological examination.</td>
</tr>
</tbody>
</table>

\textsuperscript{3} * A non-human-derived implantable foreign body (e.g. prosthetic heart valve, non-human vascular graft, mechanical heart, hip prosthesis) that is permanently placed in a patient during surgery.
The definition and measurement of surgical wound infection

BOX 7 Definition of surgical wound infection: CDC, 1992

Definitions of surgical site infection (SSI)
For surveillance classification purposes, SSIs are divided into incisional SSIs and organ/space SSIs. Incisional space SSIs are further classified as involving only the skin and subcutaneous tissue (superficial incisional SSIs) or involving deep soft tissues (e.g. fascial and muscle layers) of the incision (deep incisional SSIs). Organ/space SSIs involve any part of the anatomy (organs or spaces) other than the incision opened or manipulated during the operative procedure.

Superficial incisional SSI
Superficial incisional SSIs must meet the following criteria: infection occurs within 30 days after the operative procedure and involves only skin or subcutaneous tissue of the incision, and at least one of the following is present:
- Purulent drainage from the superficial incision.
- Organisms isolated from an aseptically obtained culture of fluid or tissue from the superficial incision.
- At least one of the following signs or symptoms of infection: pain or tenderness, localised swelling, redness or heat and superficial incision is deliberately opened by a surgeon, unless culture of incision is negative.
- Diagnosis of superficial incisional SSI by the surgeon or attending physician.

The following are not reported as superficial incisional SSIs:
- Stitch abscess (minimal inflammation and discharge confined to the points of suture penetration)
- Infection of an episiotomy or a neonate’s circumcision site* (episiotomy and circumcision are not considered NNIS operative procedures)
- Infected burn wound*
- Incisional SSI that extends into the fascial and muscle layers (see deep incisional SSI).

Deep incisional SSI
Deep incisional SSIs must meet the following criteria: infection occurs within 30 days after the operative procedure if no implant† is left in place or within 1 year if an implant is left in place and the infection appears to be related to the operative procedure and infection involves deep soft tissues (e.g. fascial and muscle layers) of the incision and at least one of the following is present:
- Purulent drainage from the deep incision but not from the organ/space component of the surgical site.
- A deep incision spontaneously dehisces or is deliberately opened by a surgeon when the patient has at least one of the following signs or symptoms: fever (> 38°C), localised pain, or tenderness, unless culture of the incision is negative.
- An abscess or other evidence of infection involving the deep incision is found on direct examination, during reoperation, or by histopathological or radiological examination.
- Diagnosis of a deep incisional SSI by a surgeon or attending physician.

Organ/space SSI
An organ/space SSI involves any part of the anatomy (e.g. organs or spaces), other than the incision, opened or manipulated during the operative procedure. Specific sites are assigned to organ/space SSIs to identify the location of the infection. Table 1 [in the original definition, see page 17 of this report] lists the specific sites that must be used to differentiate organ/space SSIs. An example is appendectomy with subsequent subdiaphragmatic abscess, which would be reported as an organ/space SSI at the intra-abdominal site.

Organ/space SSIs must meet the following criteria: infection occurs within 30 days after the operative procedure if no implant is left in place or within 1 year if an implant is left in place and the infection appears to be related to the operative procedure and infection involves any part of the anatomy (e.g. organs or spaces) other than the incision opened or manipulated during the operative procedures, and at least one of the following is present:
- Purulent drainage from a drain that is placed through a stab wound‡ into the organ/space.
- Organisms isolated from an aseptically obtained culture of fluid or tissue in the organ/space.
- An abscess or other evidence of infection involving the organ/space on direct examination, during reoperation, or by histopathological or radiological examination.
- Diagnosis of an organ/space SSI by a surgeon or attending physician.

* Specific criteria are used for infected episiotomy and circumcision sites and for burn wounds.
† An implant is defined as a non-human-derived implantable foreign body (e.g. prosthetic heart valve, non-human vascular graft, mechanical heart, hip prosthesis) that is permanently placed in a patient during operation.
‡ If the area around a stab wound becomes infected, it is not an SSI. It is considered a skin or soft tissue infection, depending on its depth.
detailed and comprehensive, with a definition for organ/space surgical site infections added.

A number of principles were given with the CDC criteria, including the statement that “a physician’s or surgeon’s diagnosis of infection derived from direct observation during surgery, endoscopic examination, or other diagnostic study, or based on clinical judgement, is an acceptable criterion for an infection, unless there is compelling evidence to the contrary”.

**Comparison between standardised definitions**

First, there are differences between the nationally proposed definitions of what constitutes a ‘surgical wound’ (*Table 4*). The CDC distinguish between superficial, deep and organ/space sites in the updated 1992 definitions. The SISG, NPS and Glenister and co-workers exclude burns, ulceration and pressure sores from their definition of ‘wound’, and the CDC exclude burns, episiotomy and neonatal circumcision. The CDC also exclude stitch abscess, defined as inflammation and discharge at the point of suture penetration.

**Timing of presentation**

There are subtle differences in the timing of presentation of surgical wound infection. The 1992 CDC definition includes infection that occurs within 30 days of surgery (without implant) or within one year of insertion of an implant if it appears related to the surgery. Glenister and co-workers define hospital acquired surgical
wound infection as that appearing at or beyond 72 hours and within 30 days without an implant or within one year if an implant is in place. The SISG definition divides presentation into early (within 30 days), intermediate (between 1 and 3 months) and late (presenting after 3 months), although no upper limit is defined for late infections. The NPS defines hospital acquired infection, including surgical wound infection, as that appearing at or beyond 48 hours after admission, but no upper limit is specified.

Individual components of surgical wound infection
The different components used to define surgical wound infection are compared in Table 5. These components are not mutually exclusive and are mostly used in combination, but illustrate the multiple factors to be considered in the assessment of surgical wounds. The Glenister and co-workers definition is the only one that uses a single criterion (the presence of purulent discharge in or exuding from a wound). The SISG/NPS definition has two criteria: either the presence of purulent discharge; or a painful, spreading erythema indicative of cellulitis. None of the standard definitions specify that positive culture of purulent discharge is a mandatory prerequisite for infection. The CDC accept the positive culture of any fluid or tissue from the superficial incision as an infection, although purulent discharge in itself fulfills the criteria. Table 5 lists the ‘stand alone’ criteria in bold face type, whereas components used in combination are non-bold face and are as described in full in Boxes 3 to 7.

Definitions identified from prospective studies
Of the 82 studies reviewed, 31 used one of the nationally proposed ‘standard’ definitions from the UK or the USA (Table 6). The CDC definitions were the most frequently referred to, being used in 29 studies, from 12 different countries. The UK definitions were used by three prospective studies, all based within the UK. Two studies, from Israel and Turkey, referred to the UK definitions in their methods section but used the CDC definitions during the study period. None of the UK studies used definitions from outwith the UK.

Other definitions of surgical wound infection
The remaining 51 studies used a variety of criteria to define surgical wound infection (Table 7). Eight of the 51 studies specified that a positive bacterial culture of purulent discharge was mandatory, and in four studies this was the only requisite for the definition of surgical wound infection. A number of studies also accepted a positive culture of organisms from a wound with drainage other than pus, similar to the CDC definition. Some studies defined infection as the presence of unspecified drainage along with signs of erythema. For example, three studies defined infection as “erythema with unspecified drainage” and one study defined it as “erythema with oedema and unspecified discharge.”

<table>
<thead>
<tr>
<th>Authors</th>
<th>Definition and exclusions</th>
</tr>
</thead>
<tbody>
<tr>
<td>CDC, 1988</td>
<td>Incisional wound: At incision site and involves skin, subcutaneous tissue or muscle located above the fascial layer. Deep wound: At operative site and involves tissues or spaces at or beneath the fascial layer.</td>
</tr>
<tr>
<td>CDC, 1992</td>
<td>Superficial incisional site: At incision site and involves only the skin and subcutaneous tissue. Exclusions: stitch abscess (minimal inflammation and discharge confined to the points of suture penetration), episiotomy or neonate’s circumcision site or burn wound. Superficial deep site: At operative site and involves deep soft tissues (e.g. fascial and muscle layers) of the incision. Organ/Space site: Organs or spaces other than the incision, opened or manipulated during the operative procedure.</td>
</tr>
<tr>
<td>SISG, NPS</td>
<td>A break in an epithelial surface, which may be surgical or accidental, drain sites included. Exclusions: burns, ulceration and pressure sores.</td>
</tr>
<tr>
<td>Glenister and co-workers, 1992</td>
<td>A break in the epithelial surface (skin or mucous membrane) and underlying tissue made by some positive act such as an accident or surgical incision. Exclusions: burns, ulcers and pressure sores.</td>
</tr>
</tbody>
</table>
The presence of drainage, purulent or otherwise, was not always a requisite for infection. A number of authors accepted the presence of erythema alone in their definition of surgical wound infection. Den Hoed and co-workers\(^8\) classed erythema or serous discharge as minor infection, and Palmer and co-workers\(^8\) accepted erythema, discharge or dehiscence. Other authors accepted different margins of extending erythema (erythema > 1 cm;\(^7\)) or erythema > 2 cm in any direction from the incision\(^8\)). Wikblad and Anderson\(^8\) defined infection as per score from a ‘redness and degree of wound healing’ grading system. Manian and Meyer\(^8\) accepted a combination of patient self-reported criteria including pain/redness, poor wound healing and persistent and intermittent fever. Four studies defined infection using a cut-off value from ASEPSIS,\(^9\) a wound grading system.\(^9\) Two studies used a score of greater than 8 in a grading system comprising erythema, induration and exudate.\(^9\)

Table 7 describes the individual components of definitions from the 51 studies. Although many studies accepted the presence of discharge with or without culture (represented by a tick in the appropriate column in the table), other criteria were also accepted, either alone or in combination. These are described in the last column of the table.
The definition and measurement of surgical wound infection

**TABLE 6** Studies that used nationally proposed standard definitions

<table>
<thead>
<tr>
<th>Author</th>
<th>Country</th>
<th>CDC, 1988(^\text{11})</th>
<th>CDC, 1992(^\text{12})</th>
<th>SISG(^\text{21}/) NPS(^\text{30})</th>
<th>Glenister and co-workers(^{28})</th>
</tr>
</thead>
<tbody>
<tr>
<td>Abramov and co-workers(^{34*})</td>
<td>Israel</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td></td>
</tr>
<tr>
<td>Anon.(^{36})</td>
<td>France</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td></td>
</tr>
<tr>
<td>Barber and co-workers(^{37})</td>
<td>USA</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td></td>
</tr>
<tr>
<td>Brown and co-workers(^{38})</td>
<td>USA</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td></td>
</tr>
<tr>
<td>Holmes and Readman(^{39})</td>
<td>UK</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td></td>
</tr>
<tr>
<td>Hopf and co-workers(^{40})</td>
<td>USA</td>
<td>✓</td>
<td>✓</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Kluytmans and co-workers(^{41})</td>
<td>Netherlands</td>
<td>✓</td>
<td>✓</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Kumarakrishnan and co-workers(^{42})</td>
<td>India</td>
<td>✓</td>
<td>✓</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Kurz and co-workers(^{43})</td>
<td>Belgium</td>
<td>✓</td>
<td>✓</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Lecuona and co-workers(^{44})</td>
<td>Spain</td>
<td>✓</td>
<td>✓</td>
<td></td>
<td></td>
</tr>
<tr>
<td>L'Ecuyer and co-workers(^{45})</td>
<td>USA</td>
<td>✓</td>
<td>✓</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Lizan-Garzia and co-workers(^{46})</td>
<td>Spain</td>
<td>✓</td>
<td>✓</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Manian and Meyer(^{47,48})</td>
<td>USA</td>
<td>✓</td>
<td>✓</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Medina-Cuadros and co-workers(^{49})</td>
<td>Spain</td>
<td>✓</td>
<td>✓</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Medina and co-workers(^{50})</td>
<td>Spain</td>
<td>✓</td>
<td>✓</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Mitchell and co-workers(^{51})</td>
<td>Australia</td>
<td>✓</td>
<td>✓</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Rantala and co-workers(^{52})</td>
<td>Finland</td>
<td>✓</td>
<td>✓</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Renz and Feliciano(^{53})</td>
<td>USA</td>
<td>✓</td>
<td>✓</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Roberts and co-workers(^{54})</td>
<td>Canada</td>
<td>✓</td>
<td>✓</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Santos and co-workers(^{55})</td>
<td>Brazil</td>
<td>✓</td>
<td>✓</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Stewart and co-workers(^{56})</td>
<td>UK</td>
<td>✓</td>
<td>✓</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Taylor and co-workers(^{57})</td>
<td>Canada</td>
<td>✓</td>
<td>✓</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Taylor and co-workers(^{58})</td>
<td>UK</td>
<td>✓</td>
<td>✓</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Taylor and co-workers(^{59})</td>
<td>Canada</td>
<td>✓</td>
<td>✓</td>
<td></td>
<td></td>
</tr>
<tr>
<td>van Griethuysen and co-workers(^{60})</td>
<td>Netherlands</td>
<td>✓</td>
<td>✓</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Vegas and co-workers(^{61})</td>
<td>Spain</td>
<td>✓</td>
<td>✓</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Velasco and co-workers(^{62})</td>
<td>Brazil</td>
<td>✓</td>
<td>✓</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Velasco and co-workers(^{63})</td>
<td>Brazil</td>
<td>✓</td>
<td>✓</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Vuorisalo and co-workers(^{64,65})</td>
<td>Finland</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td></td>
</tr>
<tr>
<td>Weiss and co-workers(^{66})</td>
<td>USA</td>
<td>✓</td>
<td>✓</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Yalcin and co-workers(^{35*})</td>
<td>Turkey</td>
<td>✓</td>
<td>✓</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

\(^*\) These studies cited references to other definitions but used the CDC definition

**Frequency of wound assessment**

Of the 82 included studies, 20 (24%) did not give details of frequency of wound assessment and details were unclear in five studies (6%). Assessment of wounds by surgical, nursing or research staff was undertaken on a daily basis in the majority of studies (55%). The remaining eight studies (10%) assessed wounds either once, twice or three times a week, and four (5%) studies were specific to day surgery.

A number of studies referred to definitions that were published before 1993. These references were traced and the definitions are included in Box 8.

**Grading of surgical wound infection**

A number of studies used grading or severity scales of surgical wound infection. These were identified


**TABLE 7 Defining criteria used by the remaining 51 studies**

<table>
<thead>
<tr>
<th>Author</th>
<th>Purulent discharge without culture</th>
<th>Culture mandatory</th>
<th>Other mandatory criteria, used alone or in combination</th>
</tr>
</thead>
<tbody>
<tr>
<td>Bellchambers and co-workers</td>
<td>–</td>
<td>–</td>
<td>ASEPSIS score &gt; 10</td>
</tr>
<tr>
<td>Bencini and co-workers</td>
<td>–</td>
<td>✔</td>
<td>Presence of at least two of the following clinical criteria: inflammation with serous discharge or purulence or presence of fibrinous debris, tenderness and warmth; and presence of leukocytes/bacteria on examination with methylene blue stain. Infection confirmed by culture in every case</td>
</tr>
<tr>
<td>Bold and co-workers</td>
<td>✔</td>
<td>–</td>
<td>Erythema and induration that required antibiotics, purulent drainage or systemic symptoms of an infection</td>
</tr>
<tr>
<td>Byrne and co-workers</td>
<td>✔</td>
<td>–</td>
<td>Or ASEPSIS score &gt; 10</td>
</tr>
<tr>
<td>den Hoed and co-workers</td>
<td>✔</td>
<td>–</td>
<td>Minor infection: erythema or serous discharge</td>
</tr>
<tr>
<td>Fanning and co-workers</td>
<td>✔</td>
<td>–</td>
<td>Major infection: purulent discharge</td>
</tr>
<tr>
<td>Fenton-Lee and co-workers</td>
<td>–</td>
<td>–</td>
<td>Purulent discharge not mandatory, will accept erythema (refer to grading scale section)</td>
</tr>
<tr>
<td>Ferraz and co-workers</td>
<td>✔</td>
<td>–</td>
<td></td>
</tr>
<tr>
<td>Gipponi and co-workers</td>
<td>✔</td>
<td>–</td>
<td></td>
</tr>
<tr>
<td>Grant and co-workers</td>
<td>✔</td>
<td>–</td>
<td>Must also have cellulitis present</td>
</tr>
<tr>
<td>Groot and Chappell</td>
<td>✔</td>
<td>–</td>
<td>Or organisms cultured from wound with seroma or erythema or both</td>
</tr>
<tr>
<td>Hakansson and co-workers</td>
<td>✔</td>
<td>–</td>
<td></td>
</tr>
<tr>
<td>Hansen and co-workers</td>
<td>✔</td>
<td>–</td>
<td></td>
</tr>
<tr>
<td>Holm and co-workers</td>
<td>✔</td>
<td>–</td>
<td>Pus with pyrexia and local tenderness</td>
</tr>
<tr>
<td>Israelsson and co-workers</td>
<td>✔</td>
<td>–</td>
<td></td>
</tr>
<tr>
<td>Israelsson and Jonsson</td>
<td>✔</td>
<td>–</td>
<td></td>
</tr>
<tr>
<td>Jewesson and co-workers</td>
<td>–</td>
<td>–</td>
<td>ASEPSIS score &gt; 21</td>
</tr>
<tr>
<td>Jewesson and co-workers</td>
<td>–</td>
<td>–</td>
<td>ASEPSIS score &gt; 21</td>
</tr>
<tr>
<td>Kingston and co-workers</td>
<td>✔</td>
<td>–</td>
<td>Major wound infection must have purulent discharge and pain and/or pyrexia</td>
</tr>
<tr>
<td>Kotisso and Asefa</td>
<td>–</td>
<td>–</td>
<td>Clinical signs of erythema, oedema and discharge</td>
</tr>
<tr>
<td>Kow and co-workers</td>
<td>✔</td>
<td>–</td>
<td>Or organisms cultured from wound with serous discharge</td>
</tr>
<tr>
<td>Kurz and co-workers</td>
<td>–</td>
<td>✔</td>
<td>ASEPSIS</td>
</tr>
<tr>
<td>Liberman and co-workers</td>
<td>–</td>
<td>–</td>
<td>Peri-incisional erythema and incisional drainage</td>
</tr>
<tr>
<td>Liem and co-workers</td>
<td>✔</td>
<td>–</td>
<td>Serious wound infection: presence of pus or sanguinopurulent discharge</td>
</tr>
<tr>
<td>Manian and Meyer</td>
<td>✔</td>
<td>–</td>
<td>(1) Must have two of three patient-reported criteria of: presence of pus or yellowish discharge; pain/redness; or poor wound healing; or persistent and intermittent fever; or (2) antibiotic treatment for surgical wound infection; or (3) physician diagnosis</td>
</tr>
<tr>
<td>Matikainen and Hiltunen</td>
<td>✔</td>
<td>–</td>
<td>Or positive culture of suppurating wound</td>
</tr>
</tbody>
</table>

*continued*
### TABLE 7 contd Defining criteria used by the remaining 51 studies

<table>
<thead>
<tr>
<th>Author</th>
<th>Purulent discharge without culture</th>
<th>Culture mandatory</th>
<th>Other mandatory criteria, used alone or in combination</th>
</tr>
</thead>
<tbody>
<tr>
<td>Milsom and co-workers</td>
<td>✔</td>
<td>–</td>
<td>Minor: erythema extending ≥ 2 cm only Major: purulent discharge with erythema</td>
</tr>
<tr>
<td>Mishriki and co-workers</td>
<td>–</td>
<td>–</td>
<td>Individual criteria of erythema, cellulitis, drainage, breakdown were recorded</td>
</tr>
<tr>
<td>Moro and co-workers</td>
<td>✔</td>
<td>–</td>
<td>Serous or non-purulent discharge with positive culture only considered infected if physical signs present</td>
</tr>
<tr>
<td>Nicols and co-workers</td>
<td>✔</td>
<td>–</td>
<td>Minor, or they did not require additional therapy or wound care</td>
</tr>
<tr>
<td>Noel and co-workers</td>
<td>–</td>
<td>–</td>
<td>Redness, discharge; purulent discharge not specified</td>
</tr>
<tr>
<td>Oertli and co-workers</td>
<td>–</td>
<td>✔</td>
<td></td>
</tr>
<tr>
<td>Palmer and co-workers</td>
<td>–</td>
<td>–</td>
<td>Erythema, discharge or dehiscence</td>
</tr>
<tr>
<td>Plattell and Hall</td>
<td>✔</td>
<td>–</td>
<td>Or serous discharge with positive culture</td>
</tr>
<tr>
<td>Poulsen and co-workers</td>
<td>✔</td>
<td>–</td>
<td>If confined to incision</td>
</tr>
<tr>
<td>Reggiori and co-workers</td>
<td>✔</td>
<td>–</td>
<td>Cellulitis with purulent exudate</td>
</tr>
<tr>
<td>Saha and co-workers</td>
<td>–</td>
<td>✔</td>
<td></td>
</tr>
<tr>
<td>Salam and co-workers</td>
<td>✔</td>
<td>–</td>
<td>Or wound discharge with positive culture</td>
</tr>
<tr>
<td>Sayed and Cade</td>
<td>–</td>
<td>✔</td>
<td>Possible wound infection if no organism isolated with ooze or redness; definite infection requires positive culture</td>
</tr>
<tr>
<td>Schein and co-workers</td>
<td>✔</td>
<td>–</td>
<td>Wound purulent requiring early removal of sutures and drainage</td>
</tr>
<tr>
<td>Serour and co-workers</td>
<td>✔</td>
<td>–</td>
<td>With redness, oedema, swelling and discharge of pus</td>
</tr>
<tr>
<td>Shirahatti and co-workers</td>
<td>✔</td>
<td>–</td>
<td>Or if redness or swelling in surrounding area</td>
</tr>
<tr>
<td>Siegman-Igra and co-workers</td>
<td>✔</td>
<td>✔</td>
<td>Must also have two of following criteria: repeated growth of same organism in culture, treatment with antibiotics or draining. Also applies to discharge other than pus</td>
</tr>
<tr>
<td>Simchen and co-workers</td>
<td>–</td>
<td>✔</td>
<td>As study by Siegman-Igra and co-workers</td>
</tr>
<tr>
<td>Slaughter and co-workers</td>
<td>✔</td>
<td>–</td>
<td>Or deemed infected by physician based on clinical judgement</td>
</tr>
<tr>
<td>Smack and co-workers</td>
<td>–</td>
<td>✔</td>
<td>Must have three symptoms: pus, erythema, tenderness and a positive culture</td>
</tr>
<tr>
<td>Stahle and co-workers</td>
<td>–</td>
<td>–</td>
<td>Signs of DSWI, mediastinitis or reoperation</td>
</tr>
<tr>
<td>Stewart and co-workers</td>
<td>✔</td>
<td>–</td>
<td>Confirmation with culture wherever possible, but not mandatory</td>
</tr>
<tr>
<td>Sturgis and co-workers</td>
<td>✔</td>
<td>–</td>
<td>Or a grading score (Jain and co-workers) greater than 8</td>
</tr>
<tr>
<td>Wikblad and Anderson</td>
<td>–</td>
<td>–</td>
<td>Evaluated on redness and degree of healing, not discharge</td>
</tr>
<tr>
<td>Wong and co-workers</td>
<td>✔</td>
<td>–</td>
<td>Presence of purulent drainage and erythema of wound edges</td>
</tr>
</tbody>
</table>

DSWI, deep sternal wound infection
from the systematic search for prospective studies, the systematic search for measurement-related studies, or from retrospective tracing of references from those identified in the original search (Table 8). The wound grading systems ranged from simple severity scales to more complex quantitative assessments of surgical wound infection. Each of the grading systems is described in more detail below, with evidence of their validity, reliability and practicality.

### The ASEPSIS grading system

The most frequently used and referred to grading system is ASEPSIS, a scale devised in 1986 by Wilson and co-workers from London. This scale (Box 9) was developed for use in a randomised controlled trial of antibiotic prophylaxis in cardiac surgery to ensure uniform reporting of all grades of wound infection found from sternal and leg wounds. ASEPSIS is an acronym for 'Additional treatment; Serous discharge; Erythema; Purulent exudate; Separation of deep tissues; Isolation of bacteria; and Stay as inpatient prolonged over 14 days'. In this initial study, a single observer examined all sternal and leg wounds daily for 5 of 7 days postoperatively for deep wound separation, serous or purulent exudate, and erythema extending 5 mm or more from the line of the incision. The proportion of the wound, as measured to the nearest 10% of its length, showing each of these features was assigned a score. Scores for purulent exudate and wound dehiscence were allocated twice the score of erythema and serous exudate, as the former were thought to be more likely to indicate wound infection. The summing of scores given over the first week was thought to reduce bias or uncertainty in the daily scores. Points were also awarded for five other objective criteria, thus not dependent upon the observer, that could be awarded at any point in the first two postoperative months. In the initial study, patients were assessed at 2 months and final scores calculated. A wound was deemed infected if the score was greater than 20. The ASEPSIS wound grading system has since been implemented in different settings and surgical specialities. A total of 19 publications related to ASEPSIS were identified from the literature search.

### The Southampton Wound Assessment Scale

The Southampton Wound Assessment Scale (Box 10) was developed by Bailey and co-workers from the University Surgical Unit in Southampton in 1992. A prospective study of elective inguinal hernia repairs was conducted over a 4-year period and all wounds were graded during inpatient stay by an experienced research nurse using the assessment chart. Community follow-up by the research nurse was undertaken at 10–14 days postoperatively and patients were subsequently...
The definition and measurement of surgical wound infection

The wound assessment scale comprises five grades, ranging in severity from 0 (normal healing) to 5 (deep or severe wound infection). Wounds were re-graded into four categories for analysis:

- normal healing
- minor complication
- wound infection – wounds graded IV or V or wounds treated with antibiotics after discharge from hospital, irrespective of the wound grading given to them by the nurse
- major haematoma – wound or scrotal haematomas requiring aspiration or evacuation.

The authors reported that the grading chart was easy to use, allowed descriptive classification and overcame the problem of subjective assessment of minor complications.

The wound grading system proposed by Coit \(142\) (Box 11) from New York was developed for use in a randomised trial of perioperative antibiotic prophylaxis in patients undergoing axillary or groin dissection. A single observer not directly involved in the care of patients assessed wounds using the scale. Wound outcome was defined as the highest grade, on a scale of 1 to 5, recorded whilst in hospital. In this study, grades 1 and 2 were classified as uncomplicated, grade 3 as minor complications, and grades 4 and 5 as major complications. The author did not report the validity, reliability or practicality of the scale. This scale was later used by Barber and co-workers \(37\) (see Box 11), from the same

<table>
<thead>
<tr>
<th>TABLE 8</th>
<th>Studies that cited other definitions and grading systems</th>
</tr>
</thead>
<tbody>
<tr>
<td>Cited by</td>
<td>Definition</td>
</tr>
<tr>
<td>Abramov and co-workers (34)</td>
<td>Platt (123,124)</td>
</tr>
<tr>
<td>Barber and co-workers (37)</td>
<td>CDC, 1992 (12)</td>
</tr>
<tr>
<td>Bellchambers and co-workers (91)</td>
<td>ASEPSIS score &gt; 10</td>
</tr>
<tr>
<td>Byrne and co-workers (92)</td>
<td>Ljundqvist (121)</td>
</tr>
<tr>
<td>den Hoed and co-workers (85)</td>
<td>As per system of Mitchell and co-workers (122)</td>
</tr>
<tr>
<td>Fenton-Lee and co-workers (86)</td>
<td>As per Cruse (436) system</td>
</tr>
<tr>
<td>Gipponi and co-workers (100)</td>
<td>Pollock (125)</td>
</tr>
<tr>
<td>Grant and co-workers (101)</td>
<td>As per grading system</td>
</tr>
<tr>
<td>Hopf and co-workers (10)</td>
<td>Simmons (433) (not obtained)</td>
</tr>
<tr>
<td>Jewesson and co-workers (93,94)</td>
<td>ASEPSIS score &gt; 21</td>
</tr>
<tr>
<td>Karran and co-workers (126)</td>
<td>As per scale</td>
</tr>
<tr>
<td>Kumarakrishnan and co-workers (42)</td>
<td>Sheridan and co-workers (434) (not obtained)</td>
</tr>
<tr>
<td>Kurz and co-workers (48)</td>
<td>Own definition</td>
</tr>
<tr>
<td>L’Ecyuer and co-workers (45)</td>
<td>CDC (version unspecified); Sawyer and Pruett (125)</td>
</tr>
<tr>
<td>Milsom and co-workers (87)</td>
<td>As per own grading system</td>
</tr>
<tr>
<td>Nichols and co-workers (10)</td>
<td>As per grading system</td>
</tr>
<tr>
<td>Poulsen and co-workers (111)</td>
<td>Simmons (433) (not obtained)</td>
</tr>
<tr>
<td>Reggiori and co-workers (112)</td>
<td>As per grading system</td>
</tr>
<tr>
<td>Salam and co-workers (80)</td>
<td>Ljundqvist (121)</td>
</tr>
<tr>
<td>Smack and co-workers (73)</td>
<td>Own definition</td>
</tr>
<tr>
<td>Smilanich and co-workers (127)</td>
<td>Altemeier and co-workers (23)</td>
</tr>
<tr>
<td>Sturgis and co-workers (96)</td>
<td>Jain and co-workers (95), grading system score &gt; 8</td>
</tr>
<tr>
<td>Wikblad and Anderson (88)</td>
<td>As per own grading system</td>
</tr>
<tr>
<td>Wilson and co-workers (128)</td>
<td>Comparison of definitions</td>
</tr>
<tr>
<td>Wilson and co-workers (129)</td>
<td>Comparison of definitions</td>
</tr>
</tbody>
</table>

reviewed at 4–6 weeks either at the outpatient department or at a domiciliary visit.
BOX 9 The ASEPSIS wound grading scale

<table>
<thead>
<tr>
<th>Wound characteristic</th>
<th>Proportion of wound infected (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>0</td>
</tr>
<tr>
<td>Serous exudate</td>
<td>0</td>
</tr>
<tr>
<td>Erythema</td>
<td>0</td>
</tr>
<tr>
<td>Purulent exudate</td>
<td>0</td>
</tr>
<tr>
<td>Separation of deep tissues</td>
<td>0</td>
</tr>
</tbody>
</table>

**Points scale for daily wound inspection**: Wounds can have a numerical score from zero (normal healing) to a possible maximum of 30, representing complete dehiscence of the wound with pus, serous exudate and erythema throughout its length. (Note: the above table states ‘serous exudate’, but in the table below the authors refer to ‘serous discharge’).

<table>
<thead>
<tr>
<th>Criterion</th>
<th>Points</th>
</tr>
</thead>
<tbody>
<tr>
<td>A</td>
<td></td>
</tr>
<tr>
<td>Antibiotics</td>
<td>10</td>
</tr>
<tr>
<td>Drainage of pus under local anaesthetic</td>
<td>5</td>
</tr>
<tr>
<td>Debridement of wound (general anaesthesia)</td>
<td>10</td>
</tr>
<tr>
<td>S</td>
<td></td>
</tr>
<tr>
<td>Serous discharge</td>
<td>Daily 0–5</td>
</tr>
<tr>
<td>E</td>
<td></td>
</tr>
<tr>
<td>Erythema</td>
<td>Daily 0–5</td>
</tr>
<tr>
<td>P</td>
<td></td>
</tr>
<tr>
<td>Purulent exudate</td>
<td>Daily 0–10</td>
</tr>
<tr>
<td>S</td>
<td></td>
</tr>
<tr>
<td>Separation of deep tissue</td>
<td>Daily 0–10</td>
</tr>
<tr>
<td>I</td>
<td></td>
</tr>
<tr>
<td>Isolation of bacteria</td>
<td>10</td>
</tr>
<tr>
<td>S</td>
<td></td>
</tr>
<tr>
<td>Stay as inpatient prolonged over 14 days</td>
<td>5</td>
</tr>
</tbody>
</table>

**Category of infection**: Total score of 0–10, satisfactory healing; 11–20, disturbance of healing; 21–30, minor wound infection; 31–40, moderate wound infection; > 40, severe wound infection.

BOX 10 The Southampton Wound Assessment Scale

<table>
<thead>
<tr>
<th>Grade</th>
<th>Appearance</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>Normal healing</td>
</tr>
<tr>
<td>I</td>
<td>Normal healing with mild bruising or haematoma</td>
</tr>
<tr>
<td>a</td>
<td>Some bruising</td>
</tr>
<tr>
<td>b</td>
<td>Considerable bruising</td>
</tr>
<tr>
<td>c</td>
<td>Mild erythema</td>
</tr>
<tr>
<td>II</td>
<td>Erythema plus other signs of inflammation</td>
</tr>
<tr>
<td>a</td>
<td>At one point</td>
</tr>
<tr>
<td>b</td>
<td>Around sutures</td>
</tr>
<tr>
<td>c</td>
<td>Along wound</td>
</tr>
<tr>
<td>d</td>
<td>Around wound</td>
</tr>
<tr>
<td>III</td>
<td>Clear or haemoserous discharge</td>
</tr>
<tr>
<td>a</td>
<td>At one point only (≤ 2 cm)</td>
</tr>
<tr>
<td>b</td>
<td>Along wound (&gt; 2 cm)</td>
</tr>
<tr>
<td>c</td>
<td>Large volume</td>
</tr>
<tr>
<td>d</td>
<td>Prolonged (&gt; 3 days)</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Grade</th>
<th>Appearance</th>
</tr>
</thead>
<tbody>
<tr>
<td>IV</td>
<td>Pus</td>
</tr>
<tr>
<td>a</td>
<td>At one point only (≤ 2 cm)</td>
</tr>
<tr>
<td>b</td>
<td>Along wound (&gt; 2 cm)</td>
</tr>
<tr>
<td>V</td>
<td>Deep or severe wound infection with or without tissue breakdown; haematoma requiring aspiration</td>
</tr>
</tbody>
</table>
The definition and measurement of surgical wound infection

Institution, in a prospective study of surgical site infection in cancer surgery. Although the 1992 CDC definition of surgical wound infection was used, the authors also classified wounds with a modified version of the Coit grading system, although no reasons were given for the modification. Wounds were graded on a scale of 1 to 4, where any surgical sites graded 3+ or 4+ were considered infected.

Fenton-Lee and co-workers from the UK used a very similar grading system (see Box 11) to assess wound complications in a study of patient acceptance and outcome of day surgery before and after service changes. Patients were followed up on days 1, 7 and 30 postoperatively by a liaison sister, who assessed wounds using the scale. The reference to the original author (Cruse, 1992) of the grading system was traced, but this reference was found to be incorrect as the scale was not found in the article.

In 1966, Karl and colleagues from Cornell University Medical College, New York, conducted a randomised trial of antibiotic prophylaxis in major surgery and classified wound infections into three grades of severity (Box 12). Modified versions of this grading system have subsequently been used

---

**BOX 11** The grading systems used by Coit, Barber and co-workers, and Fenton-Lee and co-workers

<table>
<thead>
<tr>
<th>Original grading scale: Coit, USA</th>
<th>Modified grading scale: Barber and co-workers, USA</th>
<th>Similar grading scale: Fenton-Lee and co-workers, UK</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Grade</strong></td>
<td><strong>Description</strong></td>
<td><strong>Severity</strong></td>
</tr>
<tr>
<td>1</td>
<td>Normal healing</td>
<td>Uncomplicated</td>
</tr>
<tr>
<td>2</td>
<td>Erythema within 1 cm of suture line</td>
<td>Uncomplicated</td>
</tr>
<tr>
<td>3</td>
<td>Erythema &gt; 1 cm from suture line</td>
<td>Minor</td>
</tr>
<tr>
<td>4</td>
<td>Purulent drainage, minor skin edge necrosis</td>
<td>Major</td>
</tr>
<tr>
<td>5</td>
<td>Purulent drainage, major skin edge necrosis, systemic sepsis</td>
<td>Major</td>
</tr>
</tbody>
</table>

**BOX 12** The grading systems used by Karl and co-workers, Reggiori and co-workers, and Grant and co-workers

<table>
<thead>
<tr>
<th>Original grading system: Karl and co-workers, USA</th>
<th>Modified grading system: Reggiori and co-workers, Uganda</th>
<th>Modified grading system: Grant and co-workers, USA</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Grade</strong></td>
<td><strong>Description</strong></td>
<td><strong>Grade</strong></td>
</tr>
<tr>
<td>1</td>
<td>Cellulitis with or without minimal purulent exudate</td>
<td>1</td>
</tr>
<tr>
<td>2</td>
<td>Cellulitis with moderate purulent exudate</td>
<td>2</td>
</tr>
<tr>
<td>3</td>
<td>Serous infection throughout wound or intra-abdominal abscess</td>
<td>3</td>
</tr>
</tbody>
</table>
by researchers from Uganda\textsuperscript{112} and the USA\textsuperscript{101} (see Box 12).

Jewesson and co-workers\textsuperscript{93,94} from Canada used the ASEPSIS wound grading scale (Box 13) to define surgical wound infection, citing a score of greater than 21 as a definition in two randomised controlled trials of antibiotic prophylaxis. However, they also developed a classification system with four ‘levels’ of infection. Milsom and co-workers\textsuperscript{87} from the USA, used a two-level scale (see Box 13) of minor and major infection in a randomised controlled trial of antibiotic prophylaxis in colorectal surgery. A similar scale was used by den Hoed and co-workers\textsuperscript{85} (see Box 13), from The Netherlands, in a comparative study of open versus laparoscopic cholecystectomy. This three-level system of no infection, minor and major infection was modified from one published in 1983 by Mitchell and co-workers\textsuperscript{122} (see Box 13), from the UK. Nichols and co-workers\textsuperscript{110} from the USA also used a three-level system of no, minor and major wound infection in a randomised controlled trial of antibiotic therapy for penetrating abdominal trauma.

Wikblad and Anderson\textsuperscript{88} from Sweden, compared different wound dressings in 250 patients undergoing CAGB and valve surgery. Nurses from each of three surgical units were trained to examine dressings. Assessments of dressings were made daily for 5 days, but assessment of actual wounds was undertaken only once, on day 5 post-operatively. Wounds were rated for presence of redness and extent of healing according the wound assessment protocol (Box 14).

Two final examples of wound grading systems were included in the review but were developed for evaluation of biopsy and puncture sites rather than incised wounds. Smack and co-workers\textsuperscript{73} evaluated patients undergoing biopsies in an outpatient dermatology clinic, where infection was defined by the presence of three symptoms: pus; erythema and tenderness; and a positive

<table>
<thead>
<tr>
<th>BOX 13  The grading systems used by Jewesson and co-workers,\textsuperscript{93,94} Milsom and co-workers,\textsuperscript{87} den Hoed and co-workers,\textsuperscript{85} Mitchell and co-workers\textsuperscript{122} and Nichols and co-workers\textsuperscript{110}</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Classification system; Jewesson and co-workers,\textsuperscript{93,94} Canada</strong></td>
</tr>
<tr>
<td>Class I</td>
</tr>
<tr>
<td>Class II</td>
</tr>
<tr>
<td>Class III</td>
</tr>
<tr>
<td>Class IV</td>
</tr>
<tr>
<td><strong>Milsom and co-workers,\textsuperscript{87} USA</strong></td>
</tr>
<tr>
<td>Minor</td>
</tr>
<tr>
<td>Major</td>
</tr>
<tr>
<td><strong>Modified Mitchell scale; den Hoed and co-workers,\textsuperscript{85} The Netherlands</strong></td>
</tr>
<tr>
<td>No</td>
</tr>
<tr>
<td>Minor</td>
</tr>
<tr>
<td>Major</td>
</tr>
<tr>
<td><strong>Original Mitchell scale; Mitchell and co-workers,\textsuperscript{122} UK</strong></td>
</tr>
<tr>
<td>Mild</td>
</tr>
<tr>
<td>Cellulitis</td>
</tr>
<tr>
<td><strong>Nichols and co-workers,\textsuperscript{110} USA</strong></td>
</tr>
<tr>
<td>No</td>
</tr>
<tr>
<td>Minor</td>
</tr>
<tr>
<td>Major</td>
</tr>
</tbody>
</table>
bacteriological culture. A grading system was developed to measure whether four clinical parameters of pus, erythema, tenderness and itch were not present (–), minimally present (+) or extensively present (++) . Finally, a grading system developed by Jain and co-workers was used in a randomised controlled trial of antibiotic prophylaxis in percutaneous endoscopic gastrostomy and in a subsequent trial by Sturgis and co-workers. Following gastrostomy, the peristomal area was assessed and scored daily up to 1 week for erythema (score 0 to 4), induration (score 0 to 3) and exudate (score 0 to 4). This is the only grading scale to categorise wound exudate other than just as purulent or serous; it also includes serosanguinous and sanguinous. A peristomal wound infection was diagnosed if the maximum combined score was > 8 or a purulent exudate was noted at the wound site. The percutaneous endoscopic gastrostomy procedure involves skin puncture, similar to a drain puncture site, rather than an actual wound incision.

<table>
<thead>
<tr>
<th>Redness</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Grade 0</td>
<td>No redness</td>
</tr>
<tr>
<td>Grade 1</td>
<td>Slight redness</td>
</tr>
<tr>
<td>Grade 2</td>
<td>Excessive redness</td>
</tr>
</tbody>
</table>

### Wound healing scale

<table>
<thead>
<tr>
<th>Grade 1</th>
<th>Well healed: wound edges are well together; a gap of less than 5% of the entire length of the incision is allowed, with no or slight redness</th>
</tr>
</thead>
<tbody>
<tr>
<td>Grade 2</td>
<td>Partially healed: gaps more than 5% but less than 20% of the whole length of the incision, with slight to excessive redness</td>
</tr>
<tr>
<td>Grade 3</td>
<td>Poorly healed: gaps greater than 20% of the entire length of the incision, with excessive redness</td>
</tr>
</tbody>
</table>
Chapter 4

Surveillance and monitoring of surgical wound infection

This chapter describes the published evidence on monitoring and surveillance of surgical wound infection. The literature falls broadly into four categories and this chapter is organised accordingly. First, a short overview is given of the assessment of risk of surgical wound infection and the development of wound classification and risk indices. Second, there is a substantial body of literature on hospital acquired or nosocomial infection surveillance, with examples of comprehensive national systems in the USA and Europe. Third, there are research studies specific to the diagnosis and measurement of surgical wound infection in the postdischarge environment. The 82 studies retrieved from the literature review were critiqued in terms of the duration and methodology of postdischarge surveillance and are organised according to method (e.g. studies where healthcare professionals observed surgical wounds, studies based on patient self-diagnosis of wound infection). The final section of this chapter describes the validation literature, including studies of the accuracy of case ascertainment, the accuracy of data entry methods and external system-level validation assessments. The merits, drawbacks and implications of this literature and that described in chapter 3 are brought together in chapter 5.

Quantification of risk

The risk of developing a surgical wound infection is known to be influenced by a number of factors, including the type of surgery, duration of surgery, operative technique, co-morbidity and level of contamination at the operative site. These have been described as intrinsic (e.g. susceptibility to infection due to immunosuppression) and extrinsic factors (e.g. high-risk invasive intervention; healthcare worker based or institution based). Attempts were made in the 1960s and 1970s to analyse multiple factors thought to contribute to the risk of developing surgical wound infection. The 1964 National Research Council study of the effect of ultraviolet light in the operating theatre led to the classification scheme of wound contamination according to the likelihood and degree of contamination at the time of surgery. The four categories of ‘clean’, ‘clean–contaminated’, ‘contaminated’ and ‘dirty’ were promoted by the American College of Surgeons and have since been widely accepted throughout surgery as the National Academy of Sciences/National Research Council wound classification scheme for predicting the risk of infection in surgical wounds.

<table>
<thead>
<tr>
<th>BOX 15 The National Academy of Sciences/National Research Council wound classification</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Clean wounds</strong></td>
</tr>
<tr>
<td><strong>Clean–contaminated wounds</strong></td>
</tr>
<tr>
<td><strong>Contaminated wounds</strong></td>
</tr>
<tr>
<td><strong>Dirty or infected wounds</strong></td>
</tr>
</tbody>
</table>

Estimated risk of infection

<table>
<thead>
<tr>
<th></th>
<th>1–5%</th>
<th>3–11%</th>
<th>10–17%</th>
<th>&gt; 27%</th>
</tr>
</thead>
<tbody>
<tr>
<td>Clean wounds</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Clean–contaminated wounds</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Contaminated wounds</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Dirty or infected wounds</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
In 1980, Cruse and Foord\textsuperscript{120} reported on their 10-year prospective study of almost 63,000 wounds, and reported infection rates within their own surgical centre based on the National Research Council wound classification. Surgical wound infection was defined as the discharge of pus. The wound infection rate for clean wounds was 1–2%, clean–contaminated wounds 7.7%; contaminated wounds 15.2% and dirty wounds 40%. An infection rate of clean wounds of 1–2% was deemed acceptable, but any infection rate greater than 2% for clean wounds was thought a cause for concern and investigation.

The SENIC risk index

In 1985, Haley and co-workers,\textsuperscript{147} from the USA, examined the importance of identifying individual patients at high risk of surgical wound infection. The outcome of the Study of the Efficacy of Nosocomial Infection Control (SENIC) project was the development of a simplified multivariate risk index for wound infection that could be assigned by operating room personnel. The SENIC model, based on data from 58,498 patients, identified four risk factors:

- abdominal operations
- operations lasting longer than 2 hours
- contaminated or dirty infected operation classified by the traditional wound classification system
- patients having three or more different diagnoses.

These were identified as important, independent risk factors. The authors concluded that their SENIC index predicted surgical wound infection risk about twice as accurately as the traditional wound classification system. Using this model, low-, medium- and high-risk levels of developing wound infection were identified in each of the categories of the traditional classification of wound contamination (Table 9).

The NNIS risk index

The NNIS risk index is a modification of the SENIC index and risk category is determined by allocating a point for the presence of each of the following risk factors\textsuperscript{144} (see Table 9):

- a contaminated or dirty wound class
- an ASA score of 3, 4 or 5 by the anaesthetist prior to the operative procedure
- a procedure lasting longer than $T$ hours, where $T$ is the approximate 75th percentile of the duration of surgery for that particular operative procedure.

The various operative procedures are listed on the NNIS database and, at the 75th percentile, 75% of procedures had a shorter duration of surgery and 25% had a longer duration. The NNIS risk index was designed to be used by surveillance personnel using data already available in most hospitals. In addition, the NNIS risk index was able to predict surgical wound infection procedure-specific risk. In 1991, the NNIS group suggested that the NNIS risk index was better at stratifying patients according to surgical wound infection risk than by traditional wound classification alone. According to the NNIS group, surgical wound infection rates should be stratified by risk categories before comparisons are made between institutions and surgeons or across time. It has been claimed that, thereafter, any hospital whose surgical wound infection rates in a risk category are higher than expected may have a potential problem that warrants further investigation.\textsuperscript{144}

<table>
<thead>
<tr>
<th>Risk factor</th>
<th>SENIC index</th>
<th>NNIS risk index (score 0–3)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Wound class: contaminated or dirty</td>
<td>✔</td>
<td>✔</td>
</tr>
<tr>
<td>Duration of surgery &gt; 2 h</td>
<td>✔</td>
<td>✔</td>
</tr>
<tr>
<td>Whether operation abdominal or thoracic</td>
<td>✔</td>
<td>✔</td>
</tr>
<tr>
<td>Whether patient had three or more different diagnoses</td>
<td>✔</td>
<td>✔</td>
</tr>
<tr>
<td>ASA score $\geq$ 3 prior to procedure</td>
<td>✔</td>
<td></td>
</tr>
</tbody>
</table>
Hospital-based nosocomial surveillance systems

Routine monitoring of surgical wound infection mostly falls within the bounds of hospital acquired infection surveillance. Surveillance can be performed at different levels and can be intermittent or continuous. The first stage in establishing a national programme is usually a prevalence survey, to estimate the proportion of surgical patients with wound infection at any one time. A prevalence survey quantifies the extent and potential scale of the problem and the justification for further investigation and monitoring. Different ‘levels’ of routine monitoring include:

1. targeted surveillance, where data in a defined subgroup are collected (e.g. surgical ward or department)
2. selective surveillance, where a predetermined selection of patients is monitored (e.g. all herniorrhaphy patients)
3. continuous surveillance, undertaken on an on-going basis (e.g. high-risk areas such as intensive care units)
4. intermittent surveillance, performed for a clearly specified time period.

Routine infection surveillance has been advocated as an essential component of infection control and prevention, particularly since the SENIC demonstrated that hospitals with comprehensive surveillance programmes that included feedback of surgical site infection rates to participating surgeons were associated with a decrease in subsequent infection rates. This chapter first describes the well-established American nosocomial infection system, and then expands on UK-based systems, other comprehensive national systems and individual examples of institution-based programmes.

The NNIS system, USA

The NNIS system, based at the CDC in Atlanta, was established in the USA in 1970. It is the only source of national surveillance data on nosocomial infections in the USA. This system was established in an attempt to provide a comprehensive and uniform approach to surveillance of nosocomial infection. At the beginning, hospitals adopted a policy of total or ‘hospital-wide’ surveillance, whereby all patients were monitored for nosocomial infection. By the late 1980s, separate protocols were developed for intensive care and surgical patients. Hospitals can opt to participate in the system and are then expected to develop surveillance plans using one of four standardised protocols:

- hospital-wide surveillance
- adult and paediatric intensive care unit
- high-risk nursery
- surgical patients.

The components may be used singly or simultaneously but, once selected, must be used for the minimum period of one calendar month. Participating hospitals must apply to join the surveillance programme. In 1998, a total of 276 hospitals participated in the surveillance programme. In 1999, the hospital-wide component was eliminated from the system as it was reported to be costly and time-consuming. Since 1992, the modified CDC definitions for ‘surgical site infection’ have been used. Thus, the NNIS system currently comprises targeted, selective, intermittent and continuous surveillance of surgical site infection.

The process of selected surveillance begins when infection control practitioners (ICPs) select a surgical procedure(s) from the NNIS operative list and monitor all patients undergoing that procedure for the given time period. The operative list has been modified since the 1980s and now contains over 40 surgical procedures, including laparoscopic and minimal access surgery. The use of a laparoscope (yes/no) has been collected in the NNIS system for all procedures since January 1992. Surgical infection rates are calculated and presented in various ways, including by type of operation (e.g. number of infections per 100 operations), by surgical site (e.g. number of incisional surgical site infections, deep incisional surgical site infections) and infection rates by operative procedure according to NNIS risk index category. The CDC also publish algorithms to allow hospitals to calculate surgical site infection rates with guidelines on interpretation of interhospital rates. Anonymous hospital surgical site infection rates are widely disseminated, in academic journals (e.g. American Journal of Infection Control, Infection Control & Hospital Epidemiology), on the CDC website and on individual websites of some participating hospitals.

The NINSS, England

The Nosocomial Infection National Surveillance Scheme (NINSS) was established and launched by the Department of Health and PHLS in March 1996 in response to a need for a defined programme for surveillance of infection in English hospitals. The programme is based at the Nosocomial Infection Surveillance Unit within the PHLS in London and is based on the US
surveillance system, using the 1992 CDC definitions of surgical site infection. The aim of the scheme is to develop standard methods of data collection and to provide national data for comparison. To date, two surveillance protocols have been developed:

- hospital acquired bacteraemia
- surgical site infection.

Option appraisals are underway for surveillance protocols for urinary tract infections, lower respiratory tract infections, special care baby units and intensive care units.

The NINSS surgical site infection component is selective, in that it uses a predetermined selection of patients, and intermittent, being undertaken at four time periods (January, April, July, October) for 3 consecutive months. Hospitals are required to collect data for a minimum period of 3 months and can select one or more categories for surgical procedures. Any acute hospital in England can participate, and new participants attend a surveillance workshop to learn about protocols and data collection. By January 1998, 129 hospitals had contributed data to one or both modules of hospital acquired bacteraemia and surgical site infection, and 140 hospitals had registered an interest in participating. A total of 70 hospitals had participated in surgical site infection surveillance between 1997 and 1998. At present, day-cases and endoscopic or laparoscopic procedures are excluded. The operative list comprises 12 categories: abdominal hysterectomy; bile duct, liver or pancreatic surgery; cholecystectomy without exploration of bile duct; CAGB; gastric surgery; hip prosthesis; knee prosthesis; large bowel surgery; limb amputation; open reduction of long bone fracture; small bowel surgery; and vascular surgery.

All patients undergoing the selected procedure are identified by daily reviewing of operative theatre information systems, such as theatre books, computerised records and operation lists. The study populations are prospectively monitored and visited in hospital at least three times a week from the day of surgery until discharge for a maximum of 30 days, with extended follow-up for implant surgery. Follow-up after discharge is not undertaken. Medical records and microbiology reports are reviewed to identify cases of surgical site infection. The English NINSS surgical site data collection form consists of 28 demographic, surgical and infection questions, is created and read by FORMIC, an optical mark-recognition system. Forms are returned to the NINSS on a weekly basis.

The NINSS uses the US NNIS risk index to stratify surgical wound infection rates by risk factors and to allow comparisons. Detailed reports are returned to each participating hospital to allow comparison of year-on-year rates and aggregated anonymous data from other participating hospitals. Within the NINSS, the incidence of surgical wound/site infection is reported as:

- The number of patients with one or more surgical site infections per 100 patients undergoing surgery (i.e. the risk of surgical site infection). Patients undergoing two operations in the same category are counted twice.
- The number of surgical site infections per 100 surgical operations (i.e. the ratio of surgical site infections).
- The number of surgical site infections per 1000 postoperative patient-days (i.e. the rate of surgical site infections). The number of postoperative patient-days is the number of days from the date of operation to discharge.

The NINSS aims to have reports back to individual hospitals within 2 months, although delays have been experienced due to increased demand for participation.

Wales

A surveillance system, funded by the Welsh Office, is currently being established at the PHLS, Wales. In 2000, pilot surveillance studies were conducted in five hospitals in Wales. Data on surgical procedures will be used as a denominator, and infection control teams will use the 1992 CDC definitions of surgical site infection. This surveillance system will use the FORMIC optical system. Welsh statistics will, therefore, be comparable with the English NINSS project.

Scotland

At present, Scotland does not have a national system of hospital acquired or surgical wound infection surveillance. The Scottish Office Department of Health established a multidisciplinary working group in June 1997 to develop proposals for the implementation of a national system. The Working Party findings, published in May 1999, recommended that a national framework for hospital acquired infection, including surgical site infection, should be established for the NHS in Scotland. The framework should use agreed definitions and take account of developments in England and Wales to allow direct comparison.
of data. The framework should be developed by the Scottish Centre for Infection Control and Environmental Health in conjunction with other professions and the NHS Management Executive.22

Northern Ireland
The most comprehensive reports of surgical site infection surveillance systems in Northern Ireland are from the Royal Hospitals in Belfast. Staff from this programme have worked with members from the NISS system based at the CDC in Atlanta to establish a comprehensive hospital-based surgical infection surveillance system. The 1992 CDC definitions are used. Data collection is done by theatre, ward and medical staff rather than by the infection control nurses or team. This is an alternative 'holistic' approach that encompasses staff in all aspects of data collection. The advantages of this system are that denominator data are collected on all surgical procedures. Furthermore, this team approach to data collection heightens the awareness and interest of staff. Disadvantages are that it is labour intensive as intensive education and continual encouragement are required (Smyth ETM, personal communication). Data collection forms (FORMIC questionnaire) include NNIS questions and a separate flexible section to accommodate additional questions. Forms are completed in theatre and follow patients throughout their hospital stay. They are then forwarded to the infection control team for processing and analysis. Feedback to users is conveyed both by written reports and, more recently, by interactive sessions with clinicians. A validation project is ongoing whereby an infection control nurse checks data inputted by ward and theatre staff (Smyth ETM, personal communication).

Belgium
The Institute of Hygiene and Epidemiology initiated the National Programme for Surveillance of Hospital Infections (NSIH) in Belgium in 1991.43,151–153 Specific aims of this programme were:

- to introduce and promote among hospitals the concept of surveillance of process (surgical antimicrobial prophylaxis utilisation) and outcome (surgical wound infections)
- to enable hospitals to compare their own incidence figures with those of other hospitals
- to obtain a national picture of the problem of surgical wound infection.

Following a pilot phase, the NSIH programme began in October 1992, and by 1996 was successfully implemented in two-thirds of all Belgian acute-care institutions.153 At that time, 132 hospitals participated in the scheme, with the surgical wound infection protocol being used by 57 hospitals.

There are two protocols in the Belgium NSIH system, surgical wound infections and laboratory confirmed nosocomial bloodstream infections, and either or both are implemented for a 3-month period, starting every 3 months. Participating hospitals register at least one of six classes of surgery:

- abdominal procedures
- larger selection of abdominal and general surgery
- orthopaedic surgery
- cardiovascular surgery
- gynaecology
- genitourinary surgery.

Day-case surgery is excluded. No predefined or mandatory data methods are imposed, but the protocol suggests that denominator data (procedures) be prospectively recorded by the nurse, anaesthetist or surgeon at the time of surgery in theatre in order to obtain as accurate data as possible. Data are collected on demographic variables, surgical details, ASA score, wound class and antibiotic prophylaxis. Standard data collection forms are provided but can be adapted for local use. The 1988 CDC definitions of surgical wound infection and the NNIS risk index are used.31 Data are entered at hospital level using WHOCARE software designed for surveillance of nosocomial infection by the World Health Organisation (Quality of Care and Technologies Unit of the Regional Office for Europe, Copenhagen) in collaboration with the Institute of Health and Epidemiology. This software evolved from the DANOP-DATA system software used in pre-1993 surveillance studies by Danish surgeons. Infection rates are expressed as a proportion (i.e. the number of surgical wound infections per 100 procedures performed) and as incidence density (i.e. the number of surgical wound infections per 1000 patient-observation days), thus controlling for differences in observation times. Results are returned to each participating hospital together with comparisons with national data. Validation checks on data are performed at the Institute for Health and Epidemiology, although no published details were found.

The Netherlands
In 1990, the Nosocomial Infection Committee of the National Health Council proposed a national registration system, based on data collected at a local level, for the surveillance of nosocomial
infection for hospitals in The Netherlands. The Dutch sentinel system is very similar to the Belgium surveillance programme. In 1994, Mertens and co-workers explored the potential benefit of comparing the results of the two nationwide projects and the possibility of pooling the data into one international reference data set. The data for a number of surgical categories were analysed in parallel. The study protocols are similar in the two countries, and use the 1988 CDC definitions of surgical wound infection. The Belgian hospitals use a more advanced version of the WHOCCARE software that allows the capture of the ASA score, prophylactic antibiotics and isolated organisms. The Dutch version is more limited, with wound classification, duration of surgery and antibiotic details recorded. Overall, it was reported that the patient case mix was quite different and only certain categories of surgical procedures were compared, namely appendectomy and inguinal herniorrhaphy. The comparison of international data yielded interesting differences, particularly in use of antibiotic prophylaxis (The Netherlands 3.7% versus Belgium 41.9%), although this was thought to be related to differences in risk profile, type of surgery (elective or emergency) and discharge policy. The authors stated that stratification for single risk factors was not sufficient to allow meaningful comparisons of data from different clinical settings, although the data are suitable for retrospective institutional comparison.

Other literature
Box 16 describes hospital monitoring examples from Canada, Denmark and Israel. There are many other one-off reports of prevalence surveys and surveillance reports from single institutions, including literature from Brazil, Vietnam, France, Germany and Spain. Much of the European literature on surgical wound infection is published in national nosocomial infection reports, is not available on MEDLINE and is published in languages and/or journals other than English.

Postdischarge surveillance
It was important for the review group to critique different methods of surveillance after discharge because of the increasing trend in day-case surgery and shorter duration of hospital stay. This section describes the different studies and methods of postdischarge surveillance, based on the critical appraisal of the 82 studies obtained from the literature review. Reference checking was also

<table>
<thead>
<tr>
<th>BOX 16 Examples of other surgical wound infection monitoring systems</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Canada</strong></td>
</tr>
<tr>
<td>Despite ground-breaking work in the field of surgical wound infection by Cruse and Foord in the 1970s and 1980s, Canada does not have an established national nosocomial infection surveillance programme or any system to collect standardised infection data. Although individual hospitals have established infection control mechanisms, many using the CDC definitions, much of the published data consists of single prevalence and incidence studies. The need for a standardised approach to the surveillance of surgical wound infection and centralised monitoring by Canadian institutions has been highlighted.</td>
</tr>
<tr>
<td><strong>Denmark</strong></td>
</tr>
<tr>
<td>The first prevalence studies of nosocomial infection were performed in Denmark in 1978 and 1979. In 1988, national guidelines for monitoring surgical wound infection were published by the National Centre for Hospital Hygiene. National guidelines for the registration of surgical wound infection were introduced in Denmark in 1988. Infection registration by surgical staff, rather than surveillance, is carried out in the Danish Health System. Validation of this system is described later in this chapter.</td>
</tr>
<tr>
<td><strong>Israel</strong></td>
</tr>
<tr>
<td>The Israeli Study of Surgical Infection is reported as an ongoing study where different surgical departments are surveyed for different time periods, for which every hospital within Israel is eligible. The description of the Israeli system is based on a series of articles that report a variety of data. Between 1982 and 1984 a survey of general surgery included 20 surgical departments across 11 hospitals, totalling over 5500 patients. This prevalence survey was the basis for ongoing surgical wound infection surveillance within Israeli hospitals. In the initial survey, nurse epidemiologists were trained to collect data every day using standard surveillance forms. Demographic and surgical details were precoded, but the nurse epidemiologists recorded and described wound assessments in longhand. Surgical wound infection was defined as either: (a) pus in the wound or (b) any continuous wound discharge on 2 or more days, together with at least two of the following: systemic treatment with antibiotics, local treatment such as draining, and pure culture of the same pathogen on more than one occasion. Nurse epidemiologists were present at ward and dressing rounds to record relevant clinical information. Upon discharge, surveillance forms were forwarded to a central office (Israeli Study of Surgical Infection) at the Ministry for Health, where the diagnosis of infection was made by one on-duty physician from a panel of four. Thus diagnosis was based on review of the notes recorded by the nurse epidemiologist. However, this system of secondary-level review has been severely criticised as an unnecessary layer of variability.</td>
</tr>
</tbody>
</table>
undertaken to trace articles that specifically addressed postdischarge surveillance methods or assessed the accuracy of postdischarge surveillance methods, although retrospective tracing was limited to literature published in the 1990s. This section gives an overview of the methods and duration of surveillance from the 82 included studies, and then describes in greater detail 14 studies specific to postdischarge surveillance.

Methods of postdischarge surveillance
The different methods used to detect and monitor surgical wound infection in the community setting were summarised. These included: the direct observation of wounds by medical, nursing or infection control personnel in different settings (e.g. outpatient clinic; private clinic, office, surgery) (35%); postal questionnaires to patients and/or healthcare professionals (5%); telephone interviews with patients and/or healthcare professionals (5%); and studies that used other methods or compared methods of postdischarge surveillance (9%). For the remaining studies (46%), details were either not given, unclear or postdischarge surveillance was not conducted. All the studies differed in their attempts at tracing clinic non-attendees and non-responders to questionnaires. Table 10 demonstrates the variation in duration. The majority of studies (52%) assessed wounds on or until day 30 post-operatively. Three studies assessed patients at 1 or 2 years after surgery, although this was conducted primarily to monitor other outcomes, such as hernia recurrence.

Studies of postdischarge surveillance
Fourteen articles concentrated on assessment or measurement of surgical wound infection postdischarge. Table 11 details each surveillance method, duration, study sample size and response rate. Each of these studies has been described in some detail to demonstrate the variability and lack of consistency of methods of follow-up. The studies are grouped according to whether surgical wounds were observed by healthcare professionals (direct observation) or whether diagnosis was based on patient self-report, usually by telephone or written contact. In some of the latter studies, researchers made contact with healthcare professionals in an attempt to validate patient self-diagnosis.

Observation of surgical wounds by healthcare professionals
Direct observation of surgical wounds at outpatient clinics, either by a surgeon or other healthcare professional, was the most frequent method of wound assessment after discharge from hospital. Byrne and co-workers comprehensively described their study of postdischarge surveillance in 3733 patients undergoing clean and clean–contaminated surgery. Patients’ wounds were assessed and graded using the ASEPSIS scale whilst in hospital; and infection was defined as a ‘discharge of pus’ after discharge. Patients were seen by surgeons at clinic at 6 weeks, where 70% of surgical wounds were observed directly. They achieved 99.3% follow-up using postal questionnaires to non-attendees and also employed other methods to validate patient self-reports, by contacting GPs and community nurses. The majority of infections (60%) occurred after discharge.

Ferraz and co-workers from Brazil requested patients undergoing general surgical and caesarean procedures to report at a centralised outpatient clinic on day 8 postoperatively for removal of sutures and direct observation of wounds. Telephone follow-up was not a feasible option because of the low economic status of the population and lack of ownership and access to a telephone. Over a 5-year period, return rates ranged from between 68% and 85% for general surgical patients. Timing of return was modified from day 8 to day 15 postoperatively, as the majority of infections developed after the first week. Lecuona and co-workers assessed surgical site infection diagnosed after discharge in 1103 patients undergoing general surgical procedures. Surveillance was extended to day 30 postoperatively based on a scheduled visit to the surgeon; 67% wounds were observed directly. Further attempts to follow-up non-attendees was based on a review of accident and emergency

<table>
<thead>
<tr>
<th>Maximum period of postoperative surveillance</th>
<th>No. of studies</th>
</tr>
</thead>
<tbody>
<tr>
<td>Not described in article or no follow-up after hospital discharge</td>
<td>24</td>
</tr>
<tr>
<td>1 week</td>
<td>7</td>
</tr>
<tr>
<td>2 weeks</td>
<td>4</td>
</tr>
<tr>
<td>3 weeks</td>
<td>2</td>
</tr>
<tr>
<td>1 month</td>
<td>30</td>
</tr>
<tr>
<td>6 weeks</td>
<td>10</td>
</tr>
<tr>
<td>3 months</td>
<td>2</td>
</tr>
<tr>
<td>≥ 1 year</td>
<td>3</td>
</tr>
<tr>
<td>Total</td>
<td>82</td>
</tr>
</tbody>
</table>
### TABLE 11 Methods of postdischarge surveillance

<table>
<thead>
<tr>
<th>Study</th>
<th>Country</th>
<th>Method</th>
<th>Timing</th>
<th>Sample size</th>
<th>Response rate</th>
</tr>
</thead>
<tbody>
<tr>
<td>Byrne and co-workers</td>
<td>UK</td>
<td>(1) Direct observation at outpatient clinic</td>
<td>6 weeks</td>
<td>3733</td>
<td>2426 (70%)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>(2) Questionnaire to non-attendees</td>
<td></td>
<td></td>
<td>1040 (30%)</td>
</tr>
<tr>
<td>Fanning and co-workers</td>
<td>Canada</td>
<td>(1) Patient questionnaire</td>
<td>Day 30</td>
<td>350</td>
<td>111 (32%)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>(2) Card completed by surgeon at follow-up clinic</td>
<td>4–6 weeks</td>
<td>400</td>
<td>260 (65%)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>(3) ICP questionnaire by telephone</td>
<td>Day 30</td>
<td>322</td>
<td>294 (91%)</td>
</tr>
<tr>
<td>Ferraz and co-workers</td>
<td>Brazil</td>
<td>Centralised outpatient clinic</td>
<td>Day 7 to 10</td>
<td>6604</td>
<td>68–85% over a 5-year period</td>
</tr>
<tr>
<td>Holmes and Readman</td>
<td>UK</td>
<td>(1) Completion of forms by GP</td>
<td>Day 30</td>
<td>106</td>
<td>57 (54%)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>(2) Telephone contact by infection control nurse</td>
<td></td>
<td></td>
<td>97 (92%)</td>
</tr>
<tr>
<td>Israelsson and co-workers</td>
<td>Sweden</td>
<td>Examined by surgeon</td>
<td>1 year</td>
<td>467</td>
<td>Unclear</td>
</tr>
<tr>
<td>Israelsson and co-workers</td>
<td>Sweden</td>
<td>Examined by surgeon</td>
<td>1 year</td>
<td>1023</td>
<td>Unclear</td>
</tr>
<tr>
<td>Lecuona and co-workers</td>
<td>Spain</td>
<td>Examined by surgeon</td>
<td>Day 30</td>
<td>1103</td>
<td>741 (67%)</td>
</tr>
<tr>
<td>Manian and Meyer</td>
<td>USA</td>
<td>(1) Telephone survey ≤ 3 attempts</td>
<td>Day 30</td>
<td>501</td>
<td>189 (38%)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>(2) Monthly physician questionnaire</td>
<td></td>
<td></td>
<td>Not specified</td>
</tr>
<tr>
<td>Mishriki and co-workers</td>
<td>UK</td>
<td>(1) Letter/questionnaire to pass to consulting doctor/nurse</td>
<td>Not specified</td>
<td>702</td>
<td>Not given</td>
</tr>
<tr>
<td></td>
<td></td>
<td>(2) Questionnaire survey on subsample (n = 80)</td>
<td>Not specified</td>
<td>80</td>
<td>65 (79%)</td>
</tr>
<tr>
<td>Mitchell and co-workers</td>
<td>Australia</td>
<td>(1) Mail-back questionnaire given on discharge</td>
<td>Day 28</td>
<td>1360</td>
<td>782 (58%)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>(2) Surgeon mail-back questionnaire</td>
<td>Day 28</td>
<td>1360</td>
<td>680 (50%)</td>
</tr>
<tr>
<td>Noel and co-workers</td>
<td>UK</td>
<td>(1) Patient questionnaire</td>
<td>Week 4</td>
<td>155</td>
<td>Patient 118 (76%)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>(2) GP-nurse questionnaire if infection present</td>
<td>(second mailing at week 6)</td>
<td>325</td>
<td>GP-nurse 18/28 (64%)</td>
</tr>
<tr>
<td>Roberts and co-workers</td>
<td>Canada</td>
<td>Monthly physician questionnaire</td>
<td>6 weeks</td>
<td>Unclear</td>
<td>Unclear</td>
</tr>
<tr>
<td>Sands and co-workers</td>
<td>USA</td>
<td>(1) Mailed patient questionnaires</td>
<td>Day 25 to 32</td>
<td>5572</td>
<td>1799 (33%)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>(2) Mailed surgeon questionnaires</td>
<td>4–8 weeks</td>
<td>5572</td>
<td>4420 (79%)</td>
</tr>
<tr>
<td>Santos and co-workers</td>
<td>Brazil</td>
<td>(1) Outpatient clinic</td>
<td>Day 30</td>
<td>325</td>
<td>No details on response rate</td>
</tr>
<tr>
<td></td>
<td></td>
<td>(2) Postal questionnaire</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Taylor and co-workers</td>
<td>UK</td>
<td>Diary card given to patient for GP-nurse to record symptoms and treatment</td>
<td>4–6 weeks</td>
<td>563</td>
<td>Unclear if all diary cards returned completed</td>
</tr>
</tbody>
</table>
department records, although this only increased rates of follow-up to 70%.

Santos and co-workers quantified post-discharge wound rates in 325 patients undergoing herniorrhaphy in Rio de Janeiro. Patients were examined at the outpatient surgery clinic within 30 days of surgery and infection was diagnosed using the 1992 CDC definition. At discharge, patients were also instructed to return a postal questionnaire if they experienced any wound problems. No details were given about the content of the questionnaire, the overall response rate or validation of self-reported infection. Israelson and co-workers examined patients who had undergone midline laparotomy at 1 year postoperatively for wound complications and incisional hernia. No details were given on wound assessment in the early postoperative period and whether or not wound infection rates were based on patients’ recall at 1 year postoperatively.

Use of computer-generated questionnaires

Some centres, from the USA and Canada, have reported on the use of computer-generated forms sent to surgeons for completion at the outpatient or follow-up surgical clinic. These forms are generated for each patient that the surgeon has operated on within a given time period. Manian and Meyer, first reported on the use of surgeon-specific computer-generated monthly questionnaires in 1990. This method was an attempt to improve surveillance postdischarge. Surgeons were asked to complete questionnaires at the follow-up clinic and state whether or not each patient had surgical wound infection (yes/no), and return the questionnaires on a monthly basis. Although participation in the survey was voluntary, discussions were held with surgeons and it was stressed that response was essential for improving the accuracy of infection rates and recognition of otherwise undetected outbreaks of surgical wound infection in both inpatients and outpatients. The 1988 CDC definition of surgical wound infection was prominently displayed at the beginning of the questionnaire. Details on culture, if obtained, and date of diagnosis of infection were also requested. Over a 12-month period, the overall response rate by physicians to the survey was 78%, and ranged from 57% in genitourinary surgery to 95% in vascular surgery. The majority of wound infections developed after discharge, with 14% and 71% of all infections in inpatients and outpatients, respectively, detected after discharge.

In 1997, Manian and Meyer reported their 7-year experience (1988–1995) with this method. A total of 156,977 surgical procedures were performed over this period; 43% in inpatients and 57% in ambulatory surgical patients. Overall, the annual response rate to the monthly survey was between 71% and 75% (mean 73%) over the 7-year period. Within surgical specialties, the lowest rates were found within plastic surgery and the highest within vascular surgery (98%). Of the wound infections detected, 22% were identified solely by the use of returned questionnaire (i.e. the infection would otherwise have gone undetected). The authors estimated that the time spent on the survey by the infection control department was approximately 6 hours per month, with an average of 2 hours spent on each case detected by the survey. The authors reported that the monthly computer-generated physician questionnaires complemented the hospital-based infection monitoring system.

In 1998, one Canadian tertiary care centre reported their use of computer-generated forms completed by surgeons at outpatient clinics. Surgeons were asked to complete and return infection survey forms for patients operated on in the previous 6 weeks, and asked to mark yes or no for infection (1992 CDC definition) with space provided for further details. At 6-week intervals, surgeons were asked to complete and return new forms. The authors reported a high compliance rate from surgeons with this method of postdischarge surveillance.

Patient-initiated report

Two studies examined the feasibility of telephone surveys to detect wound infection postdischarge. In 1993, Manian and Meyer explored the use of patient telephone surveys and compared the results with those obtained using the computer-generated monthly physician questionnaires. Telephone surveys were conducted on a random sample of 501 patients, with a maximum of three contact attempts made on or around day 30 postoperatively. Questions were asked about following factors specific to wound complications:

- presence of pus or yellowish discharge
- persistence of pain or redness around the incision
- poor wound healing
- persistent or intermittent fever.

Diagnosis of infection by telephone was made on the basis of at least two of the three following criteria:
• at least one of the above wound complications
• antibiotic treatment for a surgical wound infection
• physician diagnosis.

The total time spent contacting patients was equivalent to 15 minutes per successful contact, although this was only achieved in 38% of patients. Overall, this method had a poor success rate and, of those who reported having an infection, 16/18 (89%) had no evidence of infection based on the physician response to the monthly survey.

In 1995, Fanning and co-workers,98 from Canada, assessed the utility of patient telephone surveys in a randomised comparative study of three methods of postdischarge surveillance, in 1200 patients undergoing clean and clean–contaminated procedures. Surgical wound infection was defined as purulent discharge from a surgical site with or without a positive culture. Patients were randomly allocated to one of three surveillance methods (patient questionnaire, patient telephone contact, surgeon-completed diary card). The first group of patients was given a questionnaire with a stamped, addressed envelope at discharge, to be completed and returned on day 30 postoperatively. The second group of patients was contacted by an ICP by telephone on day 30 postoperatively and a standard questionnaire administered over the telephone. In the third group, surgeons were mailed a follow-up card within 2 weeks of surgery, to be returned after the clinic at 4–6 weeks postoperatively. Surgeons were provided with the case definition; information obtained from patients was evaluated by an ICP who determined whether or not the case definition was met. The results from each method are listed in Table 11. Overall, the highest rate of return was from telephone interviews (91%), but this method was labour intensive and costly. The authors questioned the reliability of patient self-report, as wounds were not directly observed, and they concluded that the most efficient method for conducting postdischarge surveillance was the clinic visit with completion of brief questionnaire cards by surgeons.

**Other methods**

Other methods used patient questionnaires or relied on patients to pass information onto healthcare professionals in the event of infection developing (Box 17). Only one study assessed the impact of late infection on community health services (Box 18).

**Validation literature**

There is a body of literature related to the accuracy, sensitivity and specificity of methods of detection and surveillance of surgical wound infection. First, there are attempts to validate case ascertainment against a gold standard (e.g. comparison of diagnosis by infection control nurses and hospital epidemiologists). Second, there are assessments of patients’ ability to self-diagnose wound infection, compared to physician diagnosis. Third, there are validation reports of data entry methods, in particular, manual versus automated data entry. Finally, there are validation studies at the level of the system and attempts to utilise existing data-collection systems (e.g. feasibility of antibiotic utilisation). Accuracy statistics from individual validation studies are listed in appendix 3.

**Validation of case ascertainment**

Cardo and co-workers,169 from one regional medical centre in the USA, conducted a prospective validation study over two time periods to determine the sensitivity and specificity of standard infection control surveillance techniques in the identification of surgical wound infection (1988 CDC definition31). Case ascertainment by three ICPs, using chart and case-note review, was compared to a gold standard of hospital epidemiologist and/or assistant hospital epidemiologist, who directly observed wounds on a daily basis. Overall, the accuracy of case ascertainment by individual ICPs was related to experience. Sensitivity levels dropped when a new ICP started, and the levels recovered thereafter. Reporting of false-positive results was not a major problem, and the authors stated that direct observation of all surgical wounds was not necessary for the accurate identification of surgical wound infection.

In a separate study at the same centre, the authors assessed the accuracy of surgical wound classification, by risk of contamination, by operating room personnel. Circulating nurses and nurse anaesthetists were taught the National Research Council wound classification system (clean, clean–contaminated, contaminated, dirty–infected) and given a short pilot period before the main study. This was compared to the gold standard (a physician observer who observed every surgical procedure), although further comparisons were also made between nurse specialists. It was found that classification errors were random and not confined to any particular category of surgery, although trauma surgery was more difficult, as multiple procedures were often performed.
In a study of patients undergoing herniorrhaphy in a day surgery unit, Holmes and Readman\textsuperscript{39} gave each patient a letter, questionnaire and prepaid envelope to be completed by their GP, or practice or district nurse, but only if the patients attended or were seen by them for any reason. Where infection was suspected (definition of purulent discharge as per Glenister and co-workers\textsuperscript{29}), the healthcare professional was requested to obtain a wound swab. Telephone contact was made with patients after 30 days and any self-report of wound infection was validated by infection control nurses, who contacted GPs. Further tracing of patients was also made by examination of the medical notes of patients who attended an outpatient clinic. Although infection developed in four patients, these were detected by telephone contact by an infection control nurse. All four patients had attended their GP, but none of the GPs returned the infection questionnaire, and infection was confirmed by contact of infection control nurses with GPs. The authors highlighted the fact that, had they only used the questionnaire method, without contact by infection control nurses, none of the wound infections would have been detected.

Mishriki and co-workers\textsuperscript{109} examined the apparent variation in rates of infection in 702 patients undergoing elective general surgery according to different components of wound infection. The salient features of wounds recorded in hospital and after discharge: erythema; cellulitis; purulent/non-purulent discharge; wound breakdown; and the doctor’s impression regarding the likelihood of infection. At discharge, all patients were given a full explanation of the study, written instructions to seek advice if symptoms developed, and also a letter with a questionnaire to pass to the consulting doctor. Wherever possible, a bacteriological swab of the wound was obtained. Symptoms of infection were reported in 62/702 (9%) patients; the most common was any reported discharge, breakdown or inflammation. Any discharge reported was also a commonly reported component. Use of clinical suspicion (or doctor’s diagnosis) rather than purulent discharge would have increased the rate of surgical wound infection from 4% to 7%. The authors also conducted a random postal survey of 80 patients; 65 questionnaires were returned (79%). Two of seven patients with symptoms in the postal survey failed to pass the questionnaire to their doctor for completion. The results from this study were, however, difficult to decipher as the symptoms were pooled from the different methods.

**Patient diary cards**

In a randomised controlled trial of antibiotic prophylaxis, Taylor and co-workers\textsuperscript{58} gave each discharged patient a diary card on which their GP or nurse was asked to record any suggestion of wound infection and details of any therapeutic intervention after discharge. Bacteriology culture swabs were also given to patients and they were asked to return them to the laboratory in the event of their wound discharging (wound infection defined as per the SISG definition: a purulent wound discharge or spreading erythema indicative of cellulitis, wound breakdown or dehiscence with clinical evidence of infection\textsuperscript{29}). All patients were reviewed at 4–6 weeks (direct observation), although details of by whom and in what setting were not given. The overall rate of infection following clean surgery was 8.9%, but had the definition been based on the presence of purulent discharge alone the overall infection rate would have been 1.2%. However, it was unclear what proportion of patients attended the clinic with completed diary cards, or whether diary cards were returned by healthcare professionals themselves.
through one incision. It was found that circulating nurses could accurately classify surgical wounds by risk of contamination and that the level of accuracy was far higher when procedures were classified into two (i.e. clean/clean-contaminated and contaminated/dirty–infected) rather than four distinct categories.

Authors from one university hospital in Brazil conducted a study to test the performance of a new method of selective chart surveillance for nosocomial infections based on risk factors identified by physicians using a risk-factor indicator form. Two trained infection control nurses reviewed the charts of surgical patients for surgical wound infection, defined as per the 1988 CDC criteria. The gold standard was chart review by two physician specialists, and validation was conducted on three separate occasions. The sensitivity of this surveillance method was 74% and the specificity was 99.7%, with an overall accuracy level of 98%.

One study in Florida compared the accuracy of classification of operative site infection by ICPs and estimated the effect of duration of surveillance experience on accuracy. The 1988 CDC definition of surgical wound infection was used throughout the Florida Consortium for Infection Control hospitals. Medical record reviewers acted as the gold standard against which case ascertainment by ICPs was compared. Sensitivity ranged from 85% to 100% and specificity from 97% to 100% across ICPs. Although there was a wide range in sensitivity values, overall the obtained values were considered satisfactory. It was found that the sensitivity of case ascertainment by ICPs improved with increasing experience, in particular it increased for ICPs with 4 or more years of experience. In a separate study, the same authors described the outcome of an epidemiological investigation of apparently over-reported (false-positive) infections in the practice of one surgeon in a community hospital. This independent external investigation found an apparent excess diagnosis of surgical site infection in laminectomy patients, this being due to incorrect diagnosis by one physician.

In the first German prevalence study of nosocomial infection, investigators were assessed on the accuracy of case ascertainment at the beginning and end of the study. Almost 15,000 patients from 72 randomly selected German hospitals were included in the study. Four investigators collected data using standardised collection forms, using the 1988 CDC definition of surgical wound infection and the NNIS risk index. The gold standard was set by two of the study supervisors experienced in infection recording. A total of 11 surgical procedures were included in the prevalence study, from traumatology (three procedures), abdominal surgery (four procedures) and gynaecology/obstetrics (four procedures). The overall sensitivity of study investigators was 89% and the specificity was 99.3%.

Although obstetric and gynaecology articles per se were excluded at the abstract review stage, one US study is included here because the authors measured the accuracy of case ascertainment. The aim of this project was to study the impact of postdischarge surveillance on the detection of nosocomial surgical site infection (1992 CDC definition) after caesarean section and vaginal delivery in almost 4500 women. Whilst in hospital, infection was diagnosed by daily review of data gathered by the infection control committee. All patients were asked to return to a postoperative clinic by day 10 to 15 for direct examination of wounds. Direct observation of all wounds was the gold standard. The sensitivity of in-hospital and postdischarge surveillance to detect surgical site infection was 16.3% and 83.7%, respectively. Although direct observation was highly sensitive, the authors reported it was time consuming and expensive, and it was not recommended for routine use. They stated it would, however, be necessary if comparisons were to be against other methods of in-hospital and postdischarge surveillance.

Validation of patient self-diagnosis

Two studies have assessed patients’ ability to self-diagnose wound infection. In 1999, Mitchell and co-workers compared rates of patient-reported and physician-reported surgical wound infection in patients undergoing major elective surgery. On discharge, each patient was given a simple questionnaire to be completed and returned after 1 month, and surgeons received a mail-back form to be completed at postoperative review. A definition of “purulent drainage present or obtained from the wound” was used. Any patient identified, either by self-report or physician report, was contacted by the research nurse for further details and confirmation. Agreement between patient self-report and physician report was measured using the κ statistic. A total of 641 forms were returned by both patients and surgeons, and both agreed that infection was present in 51 cases and absent in 565 cases (κ = 0.73). Eight surgeon-diagnosed infections
were not reported by patients and 25 patient-reported infections were not diagnosed by surgeons. However, further investigation revealed that 23 of the 25 patient-reported infections had been diagnosed by a different doctor or the infection had developed after the review by the surgeon.

Seaman and Lammers\textsuperscript{173} correlated patients’ ability to recognise wound infection with physician diagnosis in 435 patients with sutured lacerations. Patients were read a series of standard questions by a nurse, nurse practitioner, physician’s assistant, physician or medical interpreter when they presented at hospital for removal of sutures. The gold standard was judgement of wounds by a medical examiner aware of patients’ responses. A total of three nurse practitioners, 51 physicians and one physician’s assistant participated in the study. The criteria used to define wound infection were purulent discharge or tenderness, plus any two of three other signs (erythema, swelling or induration, warmth). Non-infected wounds diagnosed by medical examination were excluded from the study. Physicians were not permitted to evaluate wounds they had sutured. Wound infection was diagnosed by medical examiners in 21 patients (wound infection rate 4.8%). The nurse practitioners and physician’s assistant evaluated 70% of wounds, with the remainder being evaluated by physicians. Patients incorrectly diagnosed infection in 8% of cases and failed to recognise infection in 48% of cases. Medical examiners identified purulent drainage from 15 wounds, but eight patients (53%) did not recognise the drainage material as pus. In this population of patients with lacerated (non-surgical) wounds, the ability to self-diagnose wound infection was unreliable. This study did not attempt to measure inter-rater reliability, an important omission given the large number of participating medical assessors (n = 56). Rigorous, standardised and blinded medical assessment would have improved the strength of this study. Nevertheless, this is the only example where direct observation of wounds by both patients and medical personnel has been compared, albeit in patients with wound lacerations rather than surgically induced wounds.

Validation of data entry
In 1994, a UK study compared the accuracy of data entry methods and reported that manual data entry was 99.6% accurate compared with 99.9% for automated entry.\textsuperscript{174} Accuracy rates of less than 99.50% are thought unacceptable for automated data entry systems.\textsuperscript{174} In 1997, Smyth and co-workers,\textsuperscript{175} from the Royal Hospitals, Belfast, compared manual data entry with an automated optical scanning system. The automated system had an accuracy rate of 99.98%, with less than 0.2 errors per 1000 responses, compared with 12.4 errors per 1000 responses entered manually.

System-level validation studies
In Denmark, an external validation of the national registration system was done by conducting a concurrent prospective bedside prevalence survey.\textsuperscript{156} The routine registration system entails the recording of surgical details on a registration form by a surgeon immediately postoperatively. Any subsequent infection is detected, defined by the 1988 CDC criteria,\textsuperscript{31} and recorded by medical staff on separate documents and subsequently entered on a routine electronic hospital surveillance database. The bedside prevalence study, whereby direct observation of all wounds was performed, was considered the gold standard. The total overall sensitivity of the Danish hospital routine surveillance system was only 26% compared to case ascertainment by researchers using direct observation. The authors highlighted that many of the problems with the low sensitivity of the routine system was the lack of completion of basic registration and infection forms by surgical staff. However, they also emphasised that the Danish system is a low-cost model compared to others in Europe and the USA, which rely on specialised infection control teams.

One Dutch study assessed the feasibility of a national sentinel system using local nosocomial infection data collected by ICPs and denominator data from the Dutch National Medical Registry (Landelijke Medische Registratie (LMR)).\textsuperscript{176} The LMR database collects denominator discharge data from nearly all the hospitals in The Netherlands. This feasibility study entailed data collection by ICPs from eight hospitals for between 9 and 16 months in 1992–1993. The minimum registration period was 9 months and included general surgical, gynaecology and orthopaedic surgery. Obstetric and day-case patients were excluded. Data were collected by ICPs who visited wards twice weekly, checking charts, records and reports for signs of nosocomial infection. Patients with suspected nosocomial infection were registered and further details obtained. The 1988 CDC definition\textsuperscript{31} of surgical wound infection was used. Validation of patient findings was performed for 1 month for each study year by recording of case and denominator data and comparing data with the national LMR data source. Case finding of
patients by the local ICP was compared with a gold standard of a staff microbiologist or staff ICP. It was found that data from the LMR register could be used as a denominator, and thus would reduce the workload of infection control staff, although computer-linked data would greatly improve the process.

Examples of studies from the USA and Israel that have examined the feasibility of using routinely collected data to identify surgical site infection are given in Box 19.

<table>
<thead>
<tr>
<th>BOX 19  Use of routine data to detect surgical wound infection</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Automated medical and pharmacy records</strong></td>
</tr>
<tr>
<td>A research group from the USA designed an algorithm using statistical software to identify surgical site infections using a combination of coded elements from automated claims data, pharmacy records and automated medical records. Results were compared against the gold standard of surgical site infection retrospectively collected in a previous study. Overall, the model had a sensitivity of 74% and a specificity of 98%. The computer models created by restricting types of data sources did not perform as well when compared to when all data sources available.</td>
</tr>
<tr>
<td><strong>Antibiotic exposure</strong></td>
</tr>
<tr>
<td>Yokoe and co-workers examined the effectiveness of postoperative antibiotic exposure as a marker for surgical site infections following coronary artery bypass surgery in two samples, one selected from the USA and one selected from Israel. The 1988 CDC definitions of surgical wound infection were used in both countries. Data on almost 6000 patients were collected on a daily basis whilst patients were in hospital. Once discharged, patients were given a questionnaire over the telephone, which was repeated every 7 days until 30 days postoperatively. The gold standard was infections identified by conventional prospective surveillance methods. Postoperative antibiotic exposure with at least 9 days between the first and last day of administration had a sensitivity of 95% and 87% (USA and Israel, respectively) and a specificity of 85% and 82% (USA and Israel, respectively) for identifying surgical wound infection. The authors stated that antibiotic exposure of sufficient duration and timing was more sensitive than conventional methods in detecting surgical wound infection, but only 20% of all patients in this sample received at least 9 days of antibiotics. Quantitative antibiotic exposure thresholds for specific procedures would need to be developed to explore further this method of wound infection surveillance.</td>
</tr>
</tbody>
</table>
Chapter 5
Critical assessment of surgical wound infection

This chapter is a critique of the literature presented in chapters 3 and 4 regarding the definition, measurement and monitoring of surgical wound infection. The discussion first summarises the validity and reliability of various definitions of surgical wound infection, supplemented with evidence from formal validation studies. Second, evidence is presented of formal validation of wound infection grading and severity scales. Finally, there is a critical overview and discussion of the evidence related to the monitoring of surgical wound infection, both in hospital and after discharge.

The definition of surgical wound infection

Validity is the correct classification of infected and non-infected patients and can only be ascertained by comparing the final diagnosis to some gold standard. The diagnosis of surgical wound infection, however, is based on subjective assessment, because there is no single objective diagnostic test. The definition of surgical wound infection tends to vary around a number of properties, including redness, tenderness, swelling, warmth, purulent discharge and elevated body temperature. However, the signs and symptoms of an early wound infection can simulate an exaggerated progression of normal wound healing and a wound may show clinical signs of infection such as purulence, tenderness, warmth and erythema, but bacteria are not isolated when cultured. The converse is also possible, in that bacteria can be isolated from a healing wound that shows no clinical signs of infection. False-negative cultures can occur and, on occasion, organisms isolated from cultures may represent secondary colonisation or contamination.

The simplest definition of surgical wound infection is “purulent discharge in or exuding from a wound” as used by Glenister and co-workers and many others. This definition requires judgement of two factors: first, whether or not discharge is in or exuding from a wound; and, secondly, whether any discharge is purulent or not. However, it is restrictive, as other possible signs or symptoms of surgical wound infection, such as redness and tenderness, are excluded. Moreover, simple dichotomous definitions such as ‘presence of pus’ are practical for everyday use but fail to take severity into account. Furthermore, studies have found that the sole use of purulent discharge as a criterion may ascertain only a small proportion of wound infections and these may not necessarily be the most troublesome in the clinical setting.

A more inclusive definition is the addition of “a painful, spreading erythema indicative of cellulitis”. Examination of each component suggests that ‘painful’ will depend on subjective assessment, because there is no single objective diagnostic test. The definition of surgical wound infection tends to vary around a number of properties, including redness, tenderness, swelling, warmth, purulent discharge and elevated body temperature. However, the signs and symptoms of an early wound infection can simulate an exaggerated progression of normal wound healing and a wound may show clinical signs of infection such as purulence, tenderness, warmth and erythema, but bacteria are not isolated when cultured. The converse is also possible, in that bacteria can be isolated from a healing wound that shows no clinical signs of infection. False-negative cultures can occur and, on occasion, organisms isolated from cultures may represent secondary colonisation or contamination.

Examples of definitions with additional components include those used by the SISG and the Second UK NPS who added “infection should be considered to be present when there is fever, tenderness, oedema and an extending margin of the erythema”. The NPS definition includes a cut-off value for fever (≥ 38°C) and two additional statements:

• “or the patient is still receiving active treatment for a wound that has discharged pus”
• “the discharge of clear fluid from a wound does not indicate an infection unless accompanied by cellulitis”.

Thus, in total, the SISG accepts any of four criteria and the NPS accepts any of five criteria for the definition of surgical wound infection:

• the discharge of pus
• a painful, spreading erythema indicative of cellulitis
Critical assessment of surgical wound infection

- fever (> 38°C for NPS), tenderness, oedema and an extending margin of the erythema
- the discharge of clear fluid from a wound, accompanied by cellulitis
- the patient is still receiving active treatment for a wound that has discharged pus (NPS only).

These definitions are more sensitive and will detect higher rates of surgical wound infection than ‘presence of purulent discharge’ because of the multiple criteria. However, it should be borne in mind that there are myriad reasons for pyrexia after surgery, and these should be excluded before accepting that a pyrexia and a tender wound represent wound infection. Some of the components within the definition require assessments to be made by the patient (such as pain and/or tenderness) and at more than one point in time. Few studies have examined the ability of patients to self-diagnose infection. One Australian study\textsuperscript{51} reported good agreement between patient- and surgeon-diagnosed infection using the presence of pus as a definition. However, an American study found that patients frequently failed to recognise infection and signs of inflammation and, of those patients that had purulent discharge, half failed to recognise it as purulent.\textsuperscript{175}

A further example of multiple options within one single definition of surgical wound infection is the 1992 CDC definition,\textsuperscript{32} which lists four possible criteria for superficial site infection alone. Moreover, when dissected, the definition contains a total of six possible separate scenarios as three exist within one statement: “at least one of the following signs or symptoms of infection: pain or tenderness, localised swelling, redness, or heat – and superficial incision is deliberately opened by a surgeon, unless culture of incision is negative”. The definition, therefore, comprises:

- purulent drainage from the superficial incision, or
- organisms isolated from an aseptically obtained culture of fluid or tissue from the superficial incision, or
- pain or tenderness and superficial incision deliberately opened by a surgeon, unless culture of incision is negative, or
- localised swelling and superficial incision deliberately opened by a surgeon, unless culture of incision is negative, or
- redness or heat and superficial incision deliberately opened by a surgeon, unless culture of incision is negative, or
- diagnosis of superficial incisional surgical site infection by the surgeon or attending physician.

The CDC definition would be easier to follow if the three options within the statement given above were separated. As with the SISG and NPS definitions, some components require patient input. Wound pain and/or tenderness are present, to a variable extent, in the immediate postoperative period regardless of the presence of infection. This definition includes the term ‘heat’, which was not explicitly stated in any of the three UK definitions, and also use the term ‘localised swelling’ rather than ‘oedema’.

Although many different properties of surgical wound infection are included in the 1992 CDC definition, a number of amendments have been suggested by one senior ICP.\textsuperscript{183} First, the statement pertaining to culture of organisms should be omitted, on the basis that assessors are rarely aware of the conditions under which cultures are obtained, if cultures are obtained at all. Second, ‘pain and tenderness’ are subjective, and self-report is dependent on various patient factors, such as age and state of alertness. Gurevitch\textsuperscript{183} also argued that “or already is draining serosanguinous fluid” should be altered to read “localised swelling, redness or heat with (serosanguinous) drainage”, with the statement regarding culture removed.

The CDC are unique in their acceptance of ‘doctor’s diagnosis’ as a sole criterion for surgical wound infection. It is not accepted in any of the UK definitions reviewed, or in any grading systems. Two examples of over-reporting were found in the literature. One Canadian study\textsuperscript{184} found over-reporting by surgeons in 16% of non-infected patients when using the 1988 CDC definition. This study judged an infection as standardised if an ICP detected pus, redness or drainage associated with positive culture, and non-standardised if the ‘surgeon’s diagnosis’ was used. There was wide variation in use of ‘surgeon’s diagnosis’ within different surgical specialties, which was found to have a major impact on surgical wound infection rates.

An epidemiological investigation of high rates of surgical site infection was undertaken at a community hospital in Florida, where laminectomy patients had been operated on by one surgeon.\textsuperscript{175} The independent external investigation found an apparent excess in diagnosis of surgical site infections, or false-positive results, using the 1992 CDC definition, due to incorrect diagnosis by the surgeon. Ehrenkranz and co-workers\textsuperscript{171} proposed that use of the doctor’s diagnosis be categorised separately, as ‘presumptive’ or ‘possible’ surgical site infection and, for these patients, feedback to
surgeons should include a request for the use and documentation of objective criteria of incisional or deep infection, as approved by a hospital’s infection control committee.

Formal validation of the definition of surgical wound infection

In terms of formal validation of the definition of surgical wound infection, four articles were eligible for inclusion. Two studies evaluated the utility of objective tests to diagnose wound infection, using computed tomography (CT) and indium-111 labelled white blood cell scintigraphy. Despite being ‘objective tests’, both are open to observer variation in their interpretation. In these studies, both of which were conducted on patients undergoing cardiac surgery, clinical assessments and laboratory results were the gold standard against which the tests were compared. Although both studies demonstrated high overall sensitivities, values varied depending on the position and type of sternal prosthesis (a full description of values is given in appendix 3). Moreover, these objective tests are only applicable within cardiac surgery, for the detection of cardiothoracic complications (e.g. mediastinitis), and are not helpful in the assessment and diagnosis of surgical wound infection in the overall postsurgical population.

The remaining two studies were formal validations of different definitions of surgical wound infection. Due to the paucity of literature in this field, these studies are described in this text in detail. In 1990, Wilson and co-workers compared the definition of “discharge of pus within 28 days, irrespective of preceding serous discharge” (Cruse and Foord), with their ASEPSIS grading scale (described in chapter 3) and a three-level scale proposed by Leigh (1981). The simple Cruse and Foord definition has been used in a number of prospective studies. The Leigh grading system has three severity levels: grade 1, clinical inflammation with serous discharge but no wound breakdown; grade 2, seropurulent discharge and superficial minor wound breakdown; and grade 3, purulent discharge and major wound breakdown.

Wounds were classified using all three methods on 1029 patients. The ASEPSIS scores of greater than 10, 20, 30 and 40 points were compared to the other definitions, and the sensitivity, specificity, positive and negative predictive values were calculated by the number of wounds that required a change in management as a result of infection and that scored either more or less than the chosen threshold. There was considerable overlap of ASEPSIS scores for each of the Leigh grades. A residual wound score of over 20 points on ASEPSIS was as sensitive and significantly more specific than the Cruse and Foord definition of ‘presence of pus’, and more sensitive and specific than the Leigh and Foord grade II for prolonged stay and use of antibiotics.

Table 12, compiled using data extracted from Wilson and co-workers, demonstrates the distribution of surgical wound infection as diagnosed by ASEPSIS and by ‘purulent discharge’ (Cruse and Foord definition). The presence of purulent discharge increases as the grade of infection increases, yet it is present in only just over half of patients categorised by ASEPSIS as having moderate wound infection. This example highlights the difficulty with the inclusion of only one component, purulent discharge, as the definitive diagnosis of surgical wound infection.

Overall, the definition ‘purulent discharge’ was less sensitive than both the Leigh and Foord definition and the ASEPSIS grading system. In practical terms, however, the Cruse and Foord definition was quicker and more straightforward than ASEPSIS, which required daily assessment of wounds for scale completion.

### Table 12 Diagnosis of surgical wound infection using different criteria

<table>
<thead>
<tr>
<th>ASEPSIS score</th>
<th>ASEPSIS category</th>
<th>No. of patients classified by ASEPSIS</th>
<th>No. of patients with ‘purulent discharge’</th>
</tr>
</thead>
<tbody>
<tr>
<td>0–10</td>
<td>Satisfactory healing</td>
<td>867</td>
<td>6 (0.7%)</td>
</tr>
<tr>
<td>11–20</td>
<td>Disturbance of healing</td>
<td>74</td>
<td>13 (18%)</td>
</tr>
<tr>
<td>21–30</td>
<td>Minor wound infection</td>
<td>41</td>
<td>17 (41%)</td>
</tr>
<tr>
<td>31–40</td>
<td>Moderate wound infection</td>
<td>23</td>
<td>13 (56%)</td>
</tr>
<tr>
<td>&gt; 40</td>
<td>Severe wound infection</td>
<td>23</td>
<td>19 (83%)</td>
</tr>
<tr>
<td>Total</td>
<td></td>
<td>1028</td>
<td>68</td>
</tr>
</tbody>
</table>
The second validation study compared the 1988 CDC and NPS definitions with two grading systems (ASEPSIS, Southampton Wound Assessment Scale) over two time periods. In 1993, 325 surgical wounds were examined in 230 patients; and, in 1995, 559 surgical wounds examined in 375 patients. Data collection covered multiple surgical procedures undertaken by 32 consultant surgeons from 14 different wards. The same observer (research nurse) examined surgical wounds throughout the hospital over both time periods, and sufficient information was obtained from each patient to allow scoring of wounds by all the selected methods. In addition to inpatient assessment, to complete data collection patients were instructed to complete and return a form 1–2 months after discharge. An attempt was made to contact all patients at home and either their GP or practice nurse at 1–2 months postoperatively to ensure that all late wound infections were recorded. A total of 88% of patients were contacted directly in 1993 and 97% in 1995.

The time taken to record the information required in each of the four methods (NPS, CDC, ASEPSIS, Southampton Wound Assessment Scale) was measured on 10–20 occasions by two observers. Of 93 wounds scoring more than 20 points on the modified ASEPSIS score, 22 (24%) and 18 (19%) were not identified by the CDC and NPS definitions, respectively. The CDC and NPS definitions were less sensitive than the two scoring methods. Although the NPS appeared to detect more infections than the CDC definition (43 versus 30), the difference was not significant. Therefore, similar numbers of infected wounds were detected using the CDC and NPS definitions, and the two did not differ significantly from each other.

These two studies are the only examples of formal validation of the CDC definitions in the UK population. Further work is warranted to explore the reliability of the CDC and other definitions in order to ascertain estimates of repeatability (intra-rater reliability) and reproducibility (inter-rater reliability). There has been little investigation, within the UK at least, of the extent of variation in case ascertainment between independent trained observers, such as dedicated infection control personnel, and surgical assessors.

Formal validation of wound grading scales

ASEPSIS

The ASEPSIS wound grading system has been validated for use in patients undergoing cardiac, general and vascular surgery. In 1986, Wilson and co-workers assessed the reproducibility (inter-rater reliability) of the scale in 51 patients undergoing cardiac surgery, by calculating the coefficient of repeatability between two independent observers for sternal (n = 51) and leg (n = 34) wounds. The mean difference between observers for sternal and leg wounds was 0 and 0.1, respectively, and the coefficient of repeatability was 4.1 and 3.2, respectively. Thus 95% of differences in scores between two observers would be expected to be less than 4.1 and 3.2 for sternal and leg wounds, respectively. Although the coefficient values appear very reasonable, it depends at what point on the scale the variation occurs. The possible scores for ASEPSIS range from 0 to 150, although the four infection categories are at the lower end of the scale (0–10, 11–20, 21–30, > 40). The scale, is therefore, slightly skewed, in that a coefficient of repeatability score of 4.1 will not make a difference to a score between 40 to 150 but could make a difference at the lower end of the scale (i.e. between 0 and 40).

In 1988, Byrne and co-workers assessed the reliability (inter-rater reliability) of ASEPSIS in 100 patients undergoing general and vascular surgical procedures. Wounds were independently assessed by two observers and the mean difference was –0.1 and the correlation coefficient was 0.96. The coefficient of repeatability was 3.4 between the two observers.

As described in the definition section above, the ASEPSIS scale was compared to two definitions of surgical wound infection (Cruse and Foord, Leigh). The ASEPSIS scores of greater than 10, 20, 30 and 40 points were compared to the definitions and values calculated by the number of wounds that required a change in management as a result of infection. There was considerable overlap of ASEPSIS scores for each of the Leigh grades. An ASEPSIS wound score of over 10 points was associated with a delay in discharge from hospital, and a score of over 20 was found to be an indicator of changes in management resulting from infection. One shortcoming of the validation study was the small sample sizes, which arose due to the short duration of hospital stay.

In the later validation study by Wilson and co-workers comparing ASEPSIS and the Southampton Wound Assessment Scale and two nationally proposed definitions (1988 CDC and NPS), both grading systems were found to be more sensitive. Almost half of the wounds identified by ASEPSIS were in the minimal
disturbance of healing category. However, all methods used to detect wound infection were labour intensive, particularly in establishing the postdischarge rates.

In a review of sternal wound infection literature, Vaska described ASEPSIS as a valuable classification system for research and education purposes, but cumbersome and impractical for clinical practice. Byrne and co-workers described ASEPSIS as a ‘complex multidimensional index of wound severity’ and recommended it be standardised to a range of 0 and 1. Since ASEPSIS has been devised as an index, although this property has not been acknowledged by Wilson and co-workers or subsequent users, it should be validated on criteria by which indices are usually judged. Whilst recognising that the ASEPSIS scoring system is reliable, comparable and reproducible, Byrne and co-workers raised some questions regarding the validity of the assumptions underlying ASEPSIS. They broke down ASEPSIS into two parts. The first component is a constant, which is constructed from fixed elements of assessment (antibiotics, drainage of pus under local anaesthetic, debridement of wound under general anaesthetic, isolation of bacteria from discharge, inpatient stay > 14 days). The second component is variable, since elements are assessed daily (serous discharge, erythema, purulent exudate, separation of deep tissues). They also queried the underlying decision regarding the weighting of scores in these two parts and between elements within parts.

Their main concern, however, is that the underlying assumption that the ASEPSIS scale is linear may in fact be untrue. They report that the constant part of ASEPSIS has a maximum value of 40 (22% of the total possible score), while the variable part has a maximum value of 150 (78% of the total possible score). The variable component can have a maximum score of 30 points per day and is scored on 5 out of 7 days. Byrne and co-workers suggested that this weighting implies that the index is biased since, when there is no or very little infection (a low ASEPSIS score), the weight of the constant is greater than the variable part, while the reverse is true for severe infection. Similarly, they queried the rationale behind the different weighting of elements within parts. This weighting was not an issue for earlier scoring systems than ASEPSIS, as they only ranked questions and did not apply numerical values. Although the authors recognised that ASEPSIS is an important development in the area of postoperative wound infection, they suggested that there are areas where ASEPSIS could be improved to reduce the subjectivity of the initial evaluation of symptoms.

The Southampton Wound Assessment Scale

The Southampton Wound Assessment Scale was included in the 1990 validation study by Wilson and co-workers, where it was compared with ASEPSIS and two definitions (CDC and NPS). Of 93 wounds scoring more than 20 points (infected) on ASEPSIS, 45 (52%) were grade III or lower using the Southampton method. Grade III or lower includes normal healing (grade 1), erythema plus other signs of inflammation (grade II), and clear or haemoserous discharge (grade III). Purulent wounds are graded IV or V using the Southampton method. The 1990 study found that the Southampton Wound Assessment Scale was more sensitive than the CDC or NPS definitions at detecting surgical wound infection. There was little difference in the average time taken to complete either the Southampton scale or ASEPSIS. However, both methods were labour intensive and time consuming to complete.

One study from the University Surgical Unit in Southampton published an interim report of a trial of prophylactic antibiotic therapy in hernia surgery. The authors referred to the Southampton scale, which was developed for use in this study, stating a > 90% reproducibility between independent assessors. They also highlighted its clarity and simplicity, but did not expand on how the level of 90% reproducibility was achieved or the number of assessors used.

Other grading scales

Wikblad and Anderson assessed the reproducibility (inter-rater reliability) of their wound assessment protocol (extent of redness and wound healing). On day 5 postoperatively, a colour picture of each wound (n = 250) was taken by a university hospital photographer at a standard distance of 90 cm and angle of 90°. Two independent raters, neither of whom were aware of the experimental conditions, assessed pictures for redness and wound healing. The agreement between assessors for redness was 85%, with κ = 0.83 (very good). Agreement for the extent of wound healing was 91%, with κ = 0.81 (very good). Twelve photographs (5%) were duplicated to check repeatability (intra-rater reliability), and agreement with these ratings was 100%. However, 12 photographs (5%) is a very small sample and no description was given of the
time delay between the first and second assessments. The authors did not refer to the practicality of the instrument or give details of the timing of administration. Furthermore, assessment of photographs is not pragmatic and does not reflect the real-life situation. This study, however, is the only one that assessed repeatability (intra-rater reliability).

No further evidence of validation was found for the remaining ten grading and severity scales described in chapter 3. Several of the scales included cut-off values to aid the interpretation of a component (e.g. extent of erythema within 1 cm of the surgical wound)^[37,86,142]_. Most scales, however, included generic terms, such as ‘cellulitis’, which are open to interpretation by different assessors and therefore of questionable reliability. Subsequent differences in infection rates between studies may stem from the measurement tool or grading scale rather than be a reflection of the true rate of wound infection. Table 13 summarises the publications on the formal assessment of the definition and grading of surgical wound infection.

**Monitoring of surgical wound infection**

The assessment criteria for monitoring systems were outlined in chapter 2. In summary, these consisted of appraisal of: the use of standard definitions; the extent of coverage; the use of denominator data; the adjustment for risk; data collection and validation; and output and feedback to clinicians. The final section of this chapter is a critical overview of the evidence related to hospital and postdischarge monitoring of surgical wound infection. Other than examples of prevalence and single research studies, the bulk of the literature related to routine hospital-based monitoring of surgical wound infection falls within the realms of nosocomial or hospital-acquired infection surveillance.

**Definitions for use in routine monitoring systems**

In 1992, the US Surgical Wound Infection Task Force recommended that the CDC definitions should be adopted by all US hospitals without modification and should be used by regulatory agencies that focus on hospital infections[^33]. The most successful and extensively implemented example of hospital-based monitoring is the US NNIS study, which started in 1970. All the non-US monitoring systems reviewed in this report used either the 1988 CDC definition[^31] of surgical wound infection (incisional/deep) or the 1992 CDC definition[^32] of surgical site infection (superficial/deep/organ-space). Most of the non-US nosocomial programmes adopted the US model and conformed to some, or all, of the CDC recommended standards for data collection, use of risk indices, analysis and presentation. The CDC definition of surgical site infection has been selected for use in the recently established English NINSS and Welsh national nosocomial systems.

**Denominator data**

Most monitoring systems collect data on the population at risk of developing the adverse event, namely all patients undergoing a surgical procedure. Overall, the methods used to detect the population at risk consist of daily review of theatre information systems, computerised records and operation lists. The responsibility for the collection of denominator data, however, does vary from institution to institution, and appears to be related to whether or not a hospital has dedicated infection control personnel. For the most part, the major European nosocomial systems have adopted the American recommendations and guidelines on data collection, analysis and presentation. The English NINSS system reports wound infection in three ways: risk of infection (wound infection per 100 patients undergoing surgery), ratio of surgical infection (wound infection per 100 procedures) and rate of surgical infection (wound infections per 1000 postoperative patient-days). It is theoretically possible to compare the incidence of a postoperative adverse event (e.g. DVT) to the population as a whole, although for surgical wound infection this is not possible as the population has not undergone surgery.

**Risk adjustment**

Contamination at operation contributes to wound infection, and the National Academy of Sciences/National Research Council wound classification system[^146] relies on the surgeon’s estimate of the degree of bacterial contamination at the time of surgery[^127]. Despite various modifications, the central concept of the four-level classification system remains, although some critics argue that the presence of contamination is more significant than the degree of contamination and that the division between contaminated (class III) and dirty (class IV) is both indistinct and impractical[^127]. There is evidence from the USA to suggest that circulating theatre nurses and anaesthetic nurses could accurately classify operative procedures using the system, although accuracy levels were lower when
categorising trauma surgery, where multiple procedures were often undertaken through a single incision. The highest levels of accuracy were achieved when the classification system was reduced from four to two levels (clean/clean–contaminated surgery and contaminated/dirty surgery).

Both the SENIC and NNIS risk indices incorporate wound class and other measures of host resistance in the prediction of risk of surgical wound infection. The NNIS index is more widely used and, unlike SENIC, it can stratify risk by individual procedure. The NNIS risk index has been criticised for not including other factors, such as length of preoperative stay and other variables specific to type of surgery. However, despite criticisms, the NNIS risk index is, at present, the best method available for stratification of surgical wound infection rates, adjusting for important risk factors. The index has been widely used in European monitoring systems (e.g. Holland, Belgium) and, more recently, adopted by the English NINSS programme, which will allow intra- and inter-hospital comparisons. However, the NNIS risk index has yet to be validated in the UK patient and hospital setting, particularly with regard to the application of time limits for surgical procedures conducted in this country (Smyth ETM, personal communication).

### Data collection and validation

The majority of the reviewed hospital-based monitoring systems undertook intermittent and selective rather than continuous or hospital-wide surveillance. Data collection generally consists of assessment of patients, patient charts, medical records and laboratory reports by infection control personnel. In the UK, infection control experts from the PHLS recommended that wards should be visited daily, for review of medical, nursing and microbiology reports. With regard to data collection methodology, the gold standard for case ascertainment is direct observation and examination of the surgical wound by a trained, preferably independent, professional. Although daily direct observation of surgical wounds gives the highest sensitivity and specificity of case ascertainment, it is time consuming and costly. There is good evidence to show that ward visits, and chart and case-note reviews conducted at least three times per week by trained infection control staff yield accurate levels of case ascertainment. The evidence suggests that accuracy of data collection is highly dependent on the involvement of dedicated trained staff and, although accuracy levels tend to fluctuate in initial periods of training or recruitment, they generally rise with increased experience. Ehrenkranz and co-workers reported that ICPs with 4 or more years of experience were significantly more accurate at case ascertainment than practitioners with less experience.

Examples of alternative approaches to data collection were identified, such as the Danish surveillance programme, which relies on surgical rather than infection control staff to register and record wound infection data. External validation found that basic registration forms were
incomplete and many cases were missed, with a resulting sensitivity of case ascertainment of less than 30%. This model of data collection by surgeons proved to be very low cost compared to other European and US systems, but the trade-off was that data were effectively incomplete and accuracy levels were too low to be useful. Platell and Hall argued that junior medical and surgical staff often have little incentive and understanding of the implication of accurate data collection.

Within the national Israeli infection surveillance study, nurse epidemiologists were trained to observe surgical wounds and collect data on a daily basis. Full assessments of wounds were described in longhand on forms that were later forwarded to a central office where a clinician would ascertain whether or not infection was present. This methodology has been criticised by infection control experts as unnecessary and inaccurate. An alternative approach to wound infection data collection is that by the Royal Hospitals in Belfast, whereby theatre, nursing and surgical staff are responsible for data collection, under the supervision and encouragement of the infection control team. Formal validation of this system is currently underway.

Above all, a standardised approach to methodology facilitates inter- and intra-institutional comparison of surgical wound infection rates. For example, the study of the Dutch and Belgian systems identified many differences between the two, although many were attributed to the differences in patient case mix, use of antibiotic prophylaxis, discharge protocols and overall organisation of healthcare delivery. It is likely that wound infection data from the recently established systems in England and Wales will be comparable because of rigorous attempts to standardise data collection methodology and the overall similarities in healthcare provision. In 1997–1998, a total of 70 English hospitals participated in the wound infection programme and a further 140 hospitals had expressed an interest to the PHLS. The operative list from which procedures are selected currently covers a selection of general, orthopaedic, cardiac, gynaecological and vascular surgery. To date, laparoscopic and minimally invasive procedures have not been listed, unlike the American NISS programme.

Feedback and output
It was reported in the 1970s and 1980s that feedback of surgical wound infection rates to surgical staff contributed to reductions in subsequent infection rates. In 1985, the US SENIC estimated that a feedback programme with adequate surveillance and infection control measures, under the direction of motivated personnel, could reduce surgical wound infection rates by as much as 32%. The most efficient surveillance systems, such as the NNIS system, have dedicated resources to employ infection control staff, hospital epidemiologists and/or statisticians. The net result is a comprehensive and fluent system whereby routine and non-routine data from a number of hospital sources are collected, collated, analysed and feedback to users in a timely manner. The output from the NISS system is widely disseminated, and consists of feedback to individual hospitals, annual publications and regular website updates. The English NINSS system reports an impressive turn-around time of 2 months for participating hospitals, although some delays have been experienced due to the increased demand for participation. Feedback from the programme consists of reports of individual and aggregated rates to individual hospitals, regular newsletters and, to date, a published summary protocol. The infection control team at the Royal Hospitals in Belfast conduct interactive computer sessions and presentations with individual surgical teams, a process that is labour intensive but has promoted surgical interest and motivation. In summary, there has to be balance between timely feedback to both surgeons and institutions and collation of sufficient volumes of data to permit aggregation of rates.

Surgical wound infection as a clinical outcome indicator
Within the context of feedback to users and overall output, it is important to mention the use of surgical wound infection as a clinical outcome indicator. Reports of the use of wound infection rates as an outcome, used to reflect variation in quality of care, was first proposed during the 1980s. Today, wound infection rates are used by many as an outcome indicator to reflect quality of care within surgery. For example, in 1997, the Care Evaluation Programme of the Australian Council on Healthcare Standards introduced clinical indicators for nosocomial infection measurement. Surgical wound infection was selected as an indicator because “data were readily available, meaningful, achievable and acceptable to healthcare professionals.” The timing of measurement was limited to hospital-based infection rates measured on or from day 5 postoperatively, but only a narrow spectrum of surgery, major surgery in particular, fell within this limit. As many clean and clean-contaminated
procedures were missed because of early discharge, there has since been a move in Australia, as in other countries, including the UK (NINSS), towards calculation of procedure-specific infection rates (e.g. total hip replacement, CABG), although resource limitations have restricted the number of included procedures.

**Postdischarge surveillance**

It is well recognised that hospital rates of surgical wound infection are a gross underestimation of the true picture, and infection rates rise significantly when surveillance is extended beyond discharge. \(^{51,108}\) Postdischarge surveillance is, therefore, important in epidemiological terms. Estimates of wound infections that occur or manifest themselves after discharge range from 19% to 84%. \(^{64,172,179,194}\) The majority of these patients will be managed in the community setting, with a subsequent shift in the burden of care from primary care staff. Given the proportions of surgery conducted in the outpatient setting and the general decrease in length of inpatient stay, accurate estimation of infection occurring postdischarge is important as brief hospital stays will be insufficient to detect most infectious morbidity. \(^{58}\) According to Noel and co-workers, \(^{82}\) wound infection after dirty surgical procedures is likely to develop quickly, unlike clean surgery where infection is rarely observed before day 3 or 4 postoperatively, by which time many patients have been discharged home. Furthermore, some patients may be lost to follow-up, and thus accurate quantification of wound infection rates is difficult. \(^{82}\) Community surveillance of wound infection after clean surgery has led to the detection of higher rates than those deemed acceptable (1–2%) by the National Research Council in the 1970s. For example, one study of wound infection after clean inguinal hernia repair reported an infection rate of 9% when surveillance was extended to the community setting. \(^{141}\)

The critical appraisal of the included studies highlighted the extent of variation in data collection methods postdischarge. There are no validated systems for accurately identifying surgical wound infection that manifests after patients have left hospital. Few studies are directly comparable because of the variation in the definition of infection, staffing, location and timing of surveillance. The CDC recommend that surveillance should be conducted for 30 days postoperatively, although cardiac surgeons have argued this should be extended to 6 weeks, \(^{36}\) and others have claimed that the period could be shorter as they found the majority of infections developed within 2 weeks of surgery. \(^{99,141,194}\) Many postoperative events have resolved by the time of clinic assessment, and there is a danger that they may not be recorded. The length of time to follow-up, therefore, is still an issue of debate. Despite the selection of surgical wound infection as a benchmark of postoperative outcome, the Australian Council on Healthcare Standards guidelines failed to define a standardised method for surveillance and recommend that postdischarge data were not required and should not be collected. \(^{51}\) The Australian Council on Healthcare Standards guidelines have been criticised for their reliance on in-hospital rates and subsequent underestimation of the true scale of postoperative wound infection. \(^{51,79,195}\)

**Data collection in the community**

As with hospital-based surveillance, direct observation is the gold standard, but is labour and resource intensive, particularly within the community environment. Based on our review, the most common method used was assessment at a surgical clinic and the majority of studies conducted a final assessment within or at the end of the first postoperative month. There are several examples, from the USA and Canada, of extended surveillance from the hospital to the outpatient clinic using computer-generated forms for completion by surgeons, with completion rates of almost 100% being achieved in some surgical specialties. These systems successfully linked data obtained at follow-up to the hospital-based surgical monitoring system.

Postal and telephone surveys are subject to poor response rates, but the major drawback regarding their use is the reliance on patients’ judgement about the condition of the surgical wound. The ability of patients to accurately diagnose surgical wound infection was expanded on in the first section of this chapter, but an example was described of the inability of patients to recognise discharge material as purulent. \(^{175}\) Furthermore, some definitions distinguish and exclude stitch abscess, a distinction that patients may not be able to make. Ideally, an independent trained examiner would be more objective than would a judgement made by the patient or surgeon. With telephone surveys, one of the practical difficulties is that patients are often unavailable during working hours when infection or research staff try to contact them. Appropriate methods also vary according to socio-economic group, culture and environment. For example, in one Brazilian study telephone follow-up was not feasible due to lack of access to a telephone. \(^{29}\)
Finally, there is very little evidence of the impact of surgical wound infection on community clinical and nursing workload. One UK study estimated that postoperative wound infections doubled the time per patient spent by GPs and increased nursing time per patient at least five-fold.

**Discussion**

One objective of this methodological review was to identify valid and reliable definitions of surgical wound infection. The literature review revealed a variety of definitions, ranging from the simplest, presence of pus, to more complex, multidimensional definitions, such as those proposed by the CDC. Formal assessment, in terms of content, criterion and construct validity is limited, being partly hampered by the lack of a single gold standard objective test against which to compare definitions. Validation studies are only applicable to the setting in which they were conducted, and thus the validation of a definition or grading scale on cardiac patients in the USA may not yield the same results as those obtained in a sample of patients undergoing vascular procedures in the UK. To date, formal validation has largely been conducted within the UK setting.

Content validity is the examination of the extent to which the definition covers individual components of the topic of interest. The most straightforward definition, presence of pus, excludes other signs and symptoms of infection and thus underestimates the true rate of infection. There is evidence to suggest that this definition is less sensitive and specific than multidimensional measures. The 1992 CDC modified definition is valid in terms of content, although some of the six components have been criticised by infection control professionals. There are examples of the misuse of the ‘surgeon’s diagnosis’ component having led to inflated infection rates. In theory, the converse is equally likely, whereby surgeons underestimate the true rate of infection, particularly where the event to be measured is the validation of a definition or measure from an improvement in the underlying condition. One study, however, adopted a novel approach by photographing surgical wounds and obtaining a correlation coefficient for test–retest of two factors (presence of redness, degree of healing). Estimates of reproducibility (i.e. inter-rater reliability) have been obtained for several wound grading scales (ASEPSIS, Southampton Wound Assessment Scale). Little work has been conducted on the ability of patients to self-diagnose infection.

The available evidence from Australia and the USA is conflicting, although this may be due to differences in study robustness. Within controlled clinical trials, the repeatability and reproducibility of a definition or measurement tool is paramount, particularly where the event to be measured is the main outcome of interest, as in trials of antibiotic prophylaxis.

Standardised definitions and approaches to methodology are also fundamental for accurate monitoring of surgical wound infection. There are examples of successful monitoring systems of surgical wound infection. For the most part these fall under the realms of nosocomial or hospital acquired infection surveillance. Even well-designed, rigorous research studies and monitoring systems have difficulty in establishing infection rates postdischarge, and no standardised follow-up method is used. Each institution has its own policy and standard of practice.

In summary, surgical wound infection rates are frequently used as a clinical indicator and marker of surgical performance. Postoperative wound infection is arguably an appropriate indicator as it is a potentially modifiable cause of morbidity.
but the accurate measurement and monitoring of wound infection should be based on the use of a clear, concise and standardised definition of the event. This review has highlighted the apparent variability and inconsistency of definition, measurement and monitoring of surgical wound infection.

The bulk of surgical wound infections manifest in the community environment, yet the best means of quantifying and monitoring infection post-discharge is largely undetermined. Furthermore, the impact on primary healthcare staff is considerable and, to date, largely unquantified.
Chapter 6
Definition and measurement of anastomotic leak

Anastomotic integrity is of major concern whenever an anastomosis is constructed, as breakdown is associated with significant morbidity and mortality. Many different types of gastrointestinal anastomoses are possible (e.g. oesophagogastric, oesophagojejunal, oesophagocolic, gastrojejunial, gastrocolic, hepaticocholedochojejunal, cholangioenteric, jeunojejunial, ileocolic, ileorectal, colocolic, colorectal). Anastomotic breakdown is the most important early complication after oesophageal anastomosis, with incidence rates of up to 53% reported.\(^{196}\) Anastomotic leak is also one of the most common devastating complications after pancreatic surgery, as dehiscence of the anastomosis can lead to severe and fatal consequences due to autodigestion and destruction of surrounding tissue from leaking pancreatic juice.\(^{197}\) Anastomotic dehiscence after colorectal surgery is associated with increased perioperative mortality, due to life-threatening peritonitis and sepsis, and also affects long-term outcome as leakage is thought to influence the local recurrence of carcinoma.\(^{198}\)

In general, anastomotic leak of the gastrointestinal tract is associated with increased mortality, increased morbidity and prolonged hospital stay, and can impact on the long-term functional outcome of patients. In 1999, Pickleman and co-workers\(^{13}\) stated that it was one of the most feared complications of gastrointestinal surgery.

This chapter gives an overview and discussion of the definitions and methods used to detect and measure anastomotic leak, based on the critical appraisal of 240 articles. This chapter is presented in four sections: first, there is an overview of the excluded literature; second, the included literature is presented, broadly organised according to anatomic region; third, is an overview of studies related to diagnostic accuracy and methods used to predict, detect and measure anastomotic leak; and, finally, there is a critical appraisal and discussion of the validity and reliability of the definition and measurement of anastomotic leak.

Excluded studies
A total of 240 prospective studies identified as relevant were critically appraised and details were extracted on the definition and measurement of anastomotic leak. In total, 133 articles were excluded from the review for the following reasons:

- the study design was retrospective
- the study report did not contain a definition of anastomotic leak, a description of clinical features or a description of the investigations used to detect anastomotic leak
- anastomotic leak was not discussed in the article
- the article was a meta-analyses or review and gave no definitions, clinical features or investigations from primary studies.

A large number of publications were found to be retrospective in design, despite having ‘cohort, follow-up, or prospective’ as MeSH terms. Retrospective analyses were only included if data had been collected prospectively but analysed at a later date. Studies that reported anastomotic leak rates based on case note and medical record reviews were excluded, although many were found to have a clear definition of leakage. Review articles and meta-analyses were read, appraised and reference lists were checked for further literature.

The remaining literature contained a definition of anastomotic leak and/or details of clinical and/or radiological assessment. These are presented in three sections, according to the location of gastrointestinal surgery: upper gastrointestinal surgery, hepatopancreaticobiliary surgery and lower gastrointestinal surgery.

Studies of oesophagogastric surgery
Forty studies related to oesophagogastric surgery were eligible for inclusion. Of these, only nine included a definition of anastomotic leak and five described the features used when a clinical judgement of anastomotic leak was made (Box 20).
One classification system, published by Csendes and co-workers in 1990, was referred to by three studies and was traced and included in the review (Box 21). Csendes and co-workers conducted a prospective study of 230 patients undergoing extended total gastrectomy for advanced gastric carcinoma. The classification system was based on results from radiological investigations, whereby patients with anastomotic leak or fistula of the oesophagojejunostomy were classified as one of two types: type I (subclinical) or type II (clinical). Where there was any clinical suspicion of an anastomotic fistula in the early postoperative period, two procedures were employed:

- oral ingestion of methylene blue dissolved in water and observation of the immediate appearance of blue staining through any abdominal drain
- confirmation by an immediate X-ray study with oral ingestion of barium sulphate.

A routine barium X-ray check of the anastomosis was undertaken on all patients on day 10 postoperatively. Those with subclinical leaks had radiological studies that confirmed the presence of a local fistula involving the anastomosis, with no spillage or dissemination through a fistulous tract to the pleural or abdominal cavity, or the appearance of contrast material in any abdominal drain.

### BOX 20 Definitions of anastomotic leak from studies of oesophagogastric surgery

<table>
<thead>
<tr>
<th>Authors and co-workers</th>
<th>Definition</th>
</tr>
</thead>
<tbody>
<tr>
<td>Choi and co-workers201</td>
<td>Evidence of haematoma or seroma formation at the neck wound.</td>
</tr>
<tr>
<td>Deshmune and Shinde202</td>
<td>Defined as asymptomatic small leaks detected only at radiological study or larger leaks with a perianastomotic collection, which manifested clinically.</td>
</tr>
</tbody>
</table>
| Fernandez and co-workers203,204 | Used the Csendes classification:  
  - type I – local fistula with no spillage or dissemination through a fistulous tract to the pleural or abdominal cavity, or the appearance of contrast material in any abdominal drain  
  - type II – a leakage with great dissemination or diffusion to the pleural or abdominal cavity with the appearance of contrast medium in any of the abdominal drains. |
| Isozaki and Okajima205  | Used the Csendes classification, as above, and also categorised severity as follows:  
  - Minor leakage was defined in patients in whom leakage of contrast medium was recognised roentgenologically as a fringe-like image from the anastomotic site or was limited to a small area around the anastomotic site, but in which no leakage was recognised from the drain.  
  - Major leakage was defined as visualisation of extensive intra-abdominal contrast medium roentgenologically and leakage of contrast medium from the drain, or patients in whom symptoms of peritonitis required the insertion of a new drain. |
| Nambirajan and co-workers206 | Anastomotic leaks were defined as:  
  - incidental (small radiological leak, no clinical symptoms)  
  - minor leakage (saliva in chest drain, but clinically well)  
  - major leakage (mediastinitis or abscess, pneumothorax, empyema, radiologically confirmed major oesophageal disruption). |
| Obertop and Bosscha207  | Anastomotic disruption and mediastinitis were diagnosed by clinical signs of septicaemia in all patients and leakage confirmed by surgery in all. |
| Schardey and co-workers208 | Anastomotic insufficiency of leakage was defined as a complete intestinal wall defect at the anastomotic suture line, as detected by a contrast medium study or positive colour test (indigo carmine blue). |
| Thiede and co-workers209  | Clinically indicated by retarded bowel function, fever and localised abdominal fluid. |
| Zieren and co-workers210  | Leak defined as any radiographically demonstrable extravasation of water-soluble medium at the site of anastomosis on day 7 postoperatively. |
This classification system was used by Isozaki and Okajima\textsuperscript{205} and Fernandez and co-workers.\textsuperscript{203,204} Isozaki and Okajima analysed a large prospective data set of total and distal gastrectomies and categorised leaks as minor or major based on radiographic findings. Fernandez and co-workers conducted a prospective observational study of 101 patients undergoing total gastrectomy,\textsuperscript{203} followed by a study to assess the effect of reinforcing fibrin glue on the anastomotic leak rate on 86 patients undergoing mechanical oesophageal anastomosis.\textsuperscript{204}

Nambirajan and co-workers,\textsuperscript{206} from the UK, determined the incidence of postoperative leaks after oesophageal repair, and graded leaks, assessed both clinically and radiologically, as incidental, minor or major. A routine oesophagram was performed on all patients between day 5 and 7 to assess leakage before commencing oral feeding. Obertop and co-workers,\textsuperscript{207} from The Netherlands, reported on a small series patients ($n = 10$) who underwent oesophageal resection. ‘Clinical signs of sepsicaemia’ were used to define anastomotic leak, and all suspected cases were confirmed by re-operation. No description was given of sepsicaemia. In a German multicentre randomised controlled trial, Schardey and co-workers\textsuperscript{208} defined leak as per the results from routine contrast medium studies conducted on day 7 postoperatively. Where patients were unable or too ill to cooperate, a bedside test was performed with Gastrografin or indigo carmine blue to check the anastomosis. All test results were interpreted by radiologists who were independent of the study.

Thiede and co-workers\textsuperscript{209} conducted a European clinical trial over six centres to evaluate the applicability of a biofragmentable anastomotic ring in upper and lower gastrointestinal surgery in 1360 patients. This was the largest prospective study from the included gastrointestinal literature. Patients routinely underwent contrast radiography at one centre, but the patients from other centres only had radiological examinations in the presence of symptoms of retarded bowel function, fever or localised abdominal fluid. The methods for detecting anastomotic leak differed between centres, but no reasons were given for this.

Three reviews of upper gastrointestinal surgery were retrieved and appraised.\textsuperscript{196,211,212} Urschel\textsuperscript{211} reviewed the aetiology, prevention and management of anastomotic leaks that complicate oesophagectomy with gastric reconstruction, and reported that the incidence rate of leakage varied from $0\%$ to $30\%$, with a higher rate identified in cervical than in thoracic anastomoses. Other authors have suggested that cervical anastomoses are safer than intrathoracic anastomoses, because in the former leakage usually drains through the neck wound, whereas intrathoracic leaks can cause fatal mediastinitis.\textsuperscript{199,211} In addition, a long-term advantage is that the patient does not suffer from symptomatic gastro-oesophageal reflux of gastric contents and bile, which can be a complication of intrathoracic anastomosis.\textsuperscript{215} Although the review by Urschel\textsuperscript{211} included treatment options based on clinical and radiological presentation, it did not include definitions or methods of detection from primary studies.

Bardini and co-workers\textsuperscript{196} in a review of oesophageal anastomotic leaks, suggested a definition based on four levels or grades of severity: radiological or minor, clinical or moderate, serious, and necrosis (Box 22). This classification system requires a routine radiographic check of the anastomosis.

<table>
<thead>
<tr>
<th>BOX 22 Classification of anastomotic leak: Bardini and co-workers\textsuperscript{196}</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Radiological or minor</strong></td>
</tr>
<tr>
<td><strong>Clinical or moderate</strong></td>
</tr>
<tr>
<td><strong>Serious</strong></td>
</tr>
<tr>
<td><strong>Necrosis</strong></td>
</tr>
</tbody>
</table>

In 1998, Beitler and Urshcel\textsuperscript{212} reviewed studies that compared stapled and hand-sewn oesophago-gastric anastomoses. They analysed data from four randomised trials and seven non-randomised studies to determine which operative method was superior, using anastomotic leak rate, stricture rate and mortality as outcomes. No description was given of how leak rates were defined or detected in the primary studies, or whether leaks were defined clinically, radiologically or both, despite this being one of the main outcomes of the meta-analysis. They reported that the two main factors contributing to oesophagogastrointestinal leaks were
Definition and measurement of anastomotic leak

Technical error and occult ischaemia of the mobilised gastric fundus. Although there has been much debate in the literature over whether leak rate differs according to whether anastomoses are stapled or hand-sewn, recent evidence suggests that good results are more dependent on overall technical skill, proficiency and experience rather than choice of method.\(^{196,212}\)

Table 14 gives an overview of the details extracted from the 40 studies of upper gastrointestinal surgery and includes, where reported, the signs and symptoms used in clinical assessment, the investigations undertaken and the timing of administration. Many authors were non-specific about their assessment of patients. Most studies routinely conducted postoperative radiographic contrast swallows, using water-soluble material (usually Gastrografin) rather than barium. Two studies routinely used barium or a thin barium solution,\(^{202,214}\) and one group used barium after water-soluble contrast if no leak was detected.\(^{215}\) Two studies routinely conducted endoscopy and contrast studies on all patients postoperatively.\(^{201,216}\)

The timing of administration of routine radiological tests ranged from day 3 to day 14 postoperatively. Three studies only conducted objective tests when anastomotic leak was clinically suspected.\(^{207,217,218}\) There was considerable debate in the literature over whether or not postoperative oesophagrams should be used routinely. Nambirajan and co-workers\(^{206}\) reported that radiological asymptomatic leaks did not affect patient management, whereas minor or major leaks were clinically apparent, and that routine postoperative oesophagrams were unnecessary.

Studies of hepatopancreatico-biliary surgery

Fifteen studies related to hepatopancreatico-biliary surgery were included, 13 of which included a definition of anastomotic leak (Box 23). Most definitions were based on the presence of a high amylase concentration present in drains placed at or near the operative site.

Hardy and co-workers\(^{242}\) prospectively documented their experience of bile duct anastomosis in 129 patients undergoing liver transplantation for end-stage liver disease. Biliary complications were suspected on clinical, biochemical, microbiological and biopsy evidence, although no details of clinical assessment were given. All radiological investigations (cholangiograms) were reviewed separately by two independent observers, although no details of levels of agreement between assessors were reported.

Matsusue and co-workers\(^{241}\) distinguished between peripancreatic sepsis and pancreatic fistula, based on the amylase content and level of drainage fluid, in a study of 100 consecutive patients undergoing pancreaticejejunostomy. In a randomised trial of octreotide after pancreaticoduodenectomy, Lowy and co-workers,\(^{244}\) from the USA, clearly distinguished between a clinical pancreatic leak and a biochemical pancreatic leak. This was the only study that outlined parameters for features used in the assessment of clinical leakage. They defined fever as temperature > 38°C; leucocytosis as a white blood cell count > 10,000/l; and sepsis as haemodynamic instability requiring transfer to the intensive care unit.

In 1999, Berberat and co-workers\(^{255}\) reviewed seven randomised trials that examined the efficacy of octreotide, a potent inhibitor of exocrine pancreatic secretion, in reducing the postoperative complication rate after pancreatic surgery. The authors listed definitions for 11 postoperative complications, including leakage and pancreatic fistula, but only two of the seven trials used these definitions. Leakage was defined as “from pancreatic, biliary, or intestinal anastomosis as determined by radiographic or intraoperative findings/relaparotomy” and pancreatic fistula as “if (1) the concentration of amylase and lipase in the drainage fluid were > 3 times higher than in the serum of \(n\) consecutive postoperative days and (2) a drainage volume of > 10 ml/24 h was present, a pancreatic fistula was defined. The serum and drainage fluid amylase and/or lipase concentration were determined on postoperative days 1, 3, 4, 5 and 7 and twice weekly thereafter”. No details were given of the definitions, if any, used in the remaining five controlled trials. Berberat and co-workers did, however, present and discuss each of the seven studies separately, rather than present just a secondary analysis of individual results.

Table 15 lists the details extracted from the 15 hepatobiliary studies, including the investigations performed, the definitions of or upper values for enzyme concentrations, and the timing of the administration of tests.

There is considerable variation between separate studies in the volume of drainage fluid and the upper value of amylase concentration in drain...
<table>
<thead>
<tr>
<th>Study</th>
<th>Definition</th>
<th>Description of clinical assessment</th>
<th>Investigations</th>
<th>Timing of investigations</th>
</tr>
</thead>
<tbody>
<tr>
<td>Anikin and co-workers&lt;sup&gt;213&lt;/sup&gt;</td>
<td>–</td>
<td>Water-soluble contrast</td>
<td>Day 6</td>
<td></td>
</tr>
<tr>
<td>Choi and co-workers&lt;sup&gt;201&lt;/sup&gt;</td>
<td>–</td>
<td>Haematoma/seroma formation at neck wound (checked daily)</td>
<td>Water-soluble contrast and endoscopy</td>
<td>Day 7</td>
</tr>
<tr>
<td>Craig and co-workers&lt;sup&gt;214&lt;/sup&gt;</td>
<td>–</td>
<td>Barium contrast</td>
<td>Day 5</td>
<td></td>
</tr>
<tr>
<td>Curry and co-workers&lt;sup&gt;218&lt;/sup&gt;</td>
<td>–</td>
<td>None routinely</td>
<td>–</td>
<td></td>
</tr>
<tr>
<td>Desmane and Shinde&lt;sup&gt;202&lt;/sup&gt;</td>
<td>✔</td>
<td>Barium (thin) contrast</td>
<td>Day 10</td>
<td></td>
</tr>
<tr>
<td>Fernandez and co-workers&lt;sup&gt;203&lt;/sup&gt;</td>
<td>✔</td>
<td>Unspecified contrast</td>
<td>Day 7</td>
<td></td>
</tr>
<tr>
<td>Fernandez and co-workers&lt;sup&gt;204&lt;/sup&gt;</td>
<td>✔</td>
<td>Unspecified contrast</td>
<td>Day 7</td>
<td></td>
</tr>
<tr>
<td>Goel and co-workers&lt;sup&gt;219&lt;/sup&gt;</td>
<td>–</td>
<td>Leak at neck site</td>
<td>Water-soluble contrast</td>
<td>Day 5</td>
</tr>
<tr>
<td>Gupta&lt;sup&gt;220&lt;/sup&gt;</td>
<td>–</td>
<td>Water-soluble contrast</td>
<td>Day 5</td>
<td></td>
</tr>
<tr>
<td>Gupta and co-workers&lt;sup&gt;221&lt;/sup&gt;</td>
<td>–</td>
<td>Unspecified contrast</td>
<td>Unspecified</td>
<td></td>
</tr>
<tr>
<td>Hansson and co-workers&lt;sup&gt;222&lt;/sup&gt;</td>
<td>–</td>
<td>Water-soluble contrast</td>
<td>Day 7</td>
<td></td>
</tr>
<tr>
<td>Honkoop and co-workers&lt;sup&gt;223&lt;/sup&gt;</td>
<td>–</td>
<td>Water-soluble contrast</td>
<td>Day 7 to 10</td>
<td></td>
</tr>
<tr>
<td>Isozaki and Okajima&lt;sup&gt;205&lt;/sup&gt;</td>
<td>✔</td>
<td>Water-soluble contrast</td>
<td>Day 7 to 14</td>
<td></td>
</tr>
<tr>
<td>Jacobi and co-workers&lt;sup&gt;224&lt;/sup&gt;</td>
<td>–</td>
<td>Water-soluble contrast</td>
<td>Day 7</td>
<td></td>
</tr>
<tr>
<td>Kusano and co-workers&lt;sup&gt;225&lt;/sup&gt;</td>
<td>–</td>
<td>Unspecified radiography</td>
<td>Day 3 to 12</td>
<td></td>
</tr>
<tr>
<td>Kuwano and co-workers&lt;sup&gt;226&lt;/sup&gt;</td>
<td>–</td>
<td>Purulent discharge through drain</td>
<td>Water-soluble contrast and Urografin infusion through drain</td>
<td>Day 14</td>
</tr>
<tr>
<td>Law and co-workers&lt;sup&gt;227&lt;/sup&gt;</td>
<td>–</td>
<td>Septic complications</td>
<td>Water-soluble contrast</td>
<td>Day 7</td>
</tr>
<tr>
<td>Machens and co-workers&lt;sup&gt;228&lt;/sup&gt;</td>
<td>–</td>
<td>Local inflammation or air or saliva in cervical drain bag</td>
<td>Water-soluble contrast</td>
<td>Day 7</td>
</tr>
<tr>
<td>Mansour and co-workers&lt;sup&gt;215&lt;/sup&gt;</td>
<td>–</td>
<td>Water-soluble contrast followed by barium contrast</td>
<td>Day 6 to 9</td>
<td></td>
</tr>
<tr>
<td>Nambirajan and co-workers&lt;sup&gt;206&lt;/sup&gt;</td>
<td>✔</td>
<td>Minor: saliva in chest drain, but clinically well</td>
<td>Water-soluble contrast</td>
<td>Day 5 to 7</td>
</tr>
<tr>
<td>Obertop and Bosscha&lt;sup&gt;207&lt;/sup&gt;</td>
<td>✔</td>
<td>Unspecified contrast and CT</td>
<td>When suspected</td>
<td></td>
</tr>
<tr>
<td>O’Rourke and co-workers&lt;sup&gt;229&lt;/sup&gt;</td>
<td>–</td>
<td>Water-soluble contrast</td>
<td>Day 7 to 10</td>
<td></td>
</tr>
<tr>
<td>Pol and co-workers&lt;sup&gt;230&lt;/sup&gt;</td>
<td>–</td>
<td>Water-soluble contrast</td>
<td>Week 2</td>
<td></td>
</tr>
<tr>
<td>Rodella and co-workers&lt;sup&gt;216&lt;/sup&gt;</td>
<td>–</td>
<td>Water-soluble contrast and endoscopy</td>
<td>Unclear</td>
<td></td>
</tr>
<tr>
<td>Schardey and co-workers&lt;sup&gt;208&lt;/sup&gt;</td>
<td>✔</td>
<td>Water-soluble contrast, or indigo carmine blue test</td>
<td>Day 7</td>
<td></td>
</tr>
<tr>
<td>Schilling and co-workers&lt;sup&gt;231&lt;/sup&gt;</td>
<td>–</td>
<td>Water-soluble contrast</td>
<td>Unspecified</td>
<td></td>
</tr>
<tr>
<td>Schilling and co-workers&lt;sup&gt;232&lt;/sup&gt;</td>
<td>–</td>
<td>Water-soluble contrast</td>
<td>Day 5 to 10</td>
<td></td>
</tr>
<tr>
<td>Spiliopoulos and co-workers&lt;sup&gt;233&lt;/sup&gt;</td>
<td>–</td>
<td>Water-soluble contrast</td>
<td>Within 10 days</td>
<td></td>
</tr>
</tbody>
</table>

continued
Definition and measurement of anastomotic leak

Fluid. Some studies permit up to 50 ml/day drainage output, whereas Roder and co-workers accepted 50 ml of drainage output over the total postoperative course. Lowy and co-workers accepted an upper amylase limit of 2.5 times the normal serum amylase level to define leak, while Reissman and co-workers only accepted ten times the normal plasma level as evidence of pancreaticoduodenectomy leak.

Studies of lower gastrointestinal surgery

A total of 52 lower gastrointestinal studies were eligible for inclusion, 24 of which included a definition (Box 24). Clinical features were more commonly used in the lower gastrointestinal studies than studies of upper intestinal tract surgery and hepatopancreaticobiliary surgery. The clinical features that were frequently cited included: signs of localised or generalised peritonitis; faecal discharge from the wound and/or drain; abscess; purulent discharge from drain, wound or anus; and fever. Some studies defined a clinical anastomotic leak as that requiring re-operation, while others accepted signs of leak without further surgery.

Overall, the studies of lower gastrointestinal surgery contained more clinical trials and involved larger sample sizes than did those of upper gastrointestinal and hepatopancreaticobiliary surgery. The majority of these clinical trials included definitions. The French Association for Surgical Research conducted a controlled trial of hand-sewn anastomosis versus circular stapled anastomosis in colorectal surgery, across 24 French surgical centres, with anastomotic leak as the main outcome measure. Radiographic leaks, or covert leaks detected by sodium benzoate enema, were defined as any image that did not show a perfectly regular anastomosis and uniform calibre of lumen. The clinical signs of overt leakage included faecal matter in the drain discharge or purulent discharge per anum.

Hallbook and co-workers clearly defined clinical and radiological criteria in their multicentre study that compared straight and colonic J-pouch anastomosis, with centres included from Sweden and the USA. Docherty and co-workers, on behalf of the West of Scotland and Highland Anastomosis Study, also outlined clear criteria and definitions for their multicentre randomised controlled trial of manually constructed and stapled anastomoses in colorectal surgery. Leaks were classed as clinical and/or radiological and were presented as such. The authors noted that some patients had both clinical and radiological leaks, and it was difficult to determine the relative contribution of these on long-term prognosis.

Kessler and co-workers, from the German Study Group on Colo-Rectal Carcinoma, reported operative mortality and anastomotic leak rates

<table>
<thead>
<tr>
<th>Study</th>
<th>Definition</th>
<th>Description of clinical assessment</th>
<th>Investigations</th>
<th>Timing of investigations</th>
</tr>
</thead>
<tbody>
<tr>
<td>Svanes and co-workers</td>
<td>–</td>
<td>–</td>
<td>Water-soluble contrast</td>
<td>Day 8 to 10</td>
</tr>
<tr>
<td>Swails and co-workers</td>
<td>–</td>
<td>–</td>
<td>Water-soluble contrast</td>
<td>Day 4 to 5</td>
</tr>
<tr>
<td>Thiede and co-workers</td>
<td>✔</td>
<td>Retarded bowel function, fever, localised abdominal fluid</td>
<td>Water-soluble contrast</td>
<td>Day 8</td>
</tr>
<tr>
<td>Thomas and co-workers</td>
<td>–</td>
<td>–</td>
<td>Water-soluble contrast</td>
<td>Day 8 to 12</td>
</tr>
<tr>
<td>Trentino and co-workers</td>
<td>–</td>
<td>–</td>
<td>Water-soluble contrast</td>
<td>Day 8 to 9</td>
</tr>
<tr>
<td>van Lanschot and co-workers</td>
<td>–</td>
<td>–</td>
<td>Water-soluble contrast</td>
<td>Day 7</td>
</tr>
<tr>
<td>Vigneswaran and co-workers</td>
<td>–</td>
<td>–</td>
<td>Water-soluble contrast</td>
<td>Day 7</td>
</tr>
<tr>
<td>Wu and co-workers</td>
<td>–</td>
<td>–</td>
<td>Water-soluble contrast</td>
<td>Unspecified</td>
</tr>
<tr>
<td>Zieren and co-workers</td>
<td>✔</td>
<td>–</td>
<td>Water-soluble contrast</td>
<td>Day 7</td>
</tr>
<tr>
<td>Zilling and co-workers</td>
<td>–</td>
<td>–</td>
<td>Water-soluble contrast</td>
<td>Day 4 to 7</td>
</tr>
</tbody>
</table>

* Thiede and co-workers conducted upper and lower gastrointestinal surgery, but their study is included here as the majority of procedures (63%) were upper anastomoses.
## BOX 23 Definitions of anastomotic leak from studies of hepatopancreatobiliary surgery

<table>
<thead>
<tr>
<th>Authors</th>
<th>Definition</th>
</tr>
</thead>
<tbody>
<tr>
<td>Bottger and co-workers</td>
<td>Pancreatic fistula was identified in the presence of an amylase concentration in drainage fluid in excess of 2000 units/l.</td>
</tr>
<tr>
<td>Chou and co-workers</td>
<td>A pancreatic leak or fistula was defined as persistent drainage of ≥ 50 ml/day of amylase-rich fluid for more than 2 weeks.</td>
</tr>
<tr>
<td>Cullen and co-workers</td>
<td>The recovery of fluid from the peripancreatic drains with an amylase content five times greater than normal; leakage was demonstrated radiographically or anastomotic disruption was confirmed at re-operation.</td>
</tr>
<tr>
<td>Evans and co-workers</td>
<td>The presence of a pancreatic fistula was defined by the production of &gt; 50 ml/day abdominal fluid with an amylase content of &gt; 1000 units/l (normally &lt; 300 units/l).</td>
</tr>
<tr>
<td>Gupta and co-workers</td>
<td>Anastomotic leaks were identified as bile drainage from drains placed at the area of the anastomosis.</td>
</tr>
<tr>
<td>Hamanka and Suzuki</td>
<td>Daily pancreatic juice was measured and monitored during the postoperative period until the tube spontaneously discharged; leakage based on clinical signs (peritonitis, pyrexia, sepsis).</td>
</tr>
<tr>
<td>Howard</td>
<td>Serum amylase levels were determined at 2-day intervals for the days 8–12 postoperatively. Amylase levels were determined on the same day on the peritoneal (drain) and pancreatic (tube) fluid, respectively, at least once between day 7 and 10 postoperatively.</td>
</tr>
<tr>
<td>Kayahara and co-workers</td>
<td>A diagnosis of dehiscence of the anastomosis was made in accordance with any of the following criteria: the drainage of bile or enteric juice from the drain; the detection of enteric bacteria in the drainage fluid; radiographic confirmation of dehiscence of pancreatic ductography; or an amylase level in the drainage fluid of more than three times the serum amylase level.</td>
</tr>
<tr>
<td>Lowy and co-workers</td>
<td>A clinical pancreatic anastomotic leak was defined as the drainage of amylase-rich fluid (&gt; 2.5 times the upper limit of normal for serum amylase) in association with fever (&gt; 38°C), leucocytosis (white blood cell count &gt; 10,000/1), sepsis (haemodynamic instability requiring transfer to the intensive care unit), or the need for percutaneous drainage of an amylase-rich fluid collection. A biochemical pancreatic leak was defined as an elevated level of amylase (&gt; 2.5 times the upper limit of normal for serum amylase) in the drain fluid on or after postoperative day 3 that was asymptomatic and resolved spontaneously.</td>
</tr>
<tr>
<td>Matsusue and co-workers</td>
<td>Protracted healing of pancreaticojejunostomy was divided into peripancreatic sepsis and pancreatic fistula. Peripancreatic sepsis was defined as prolonged suppurative discharge of &lt; 50 ml/day with a low amylase content (&lt; 1000 IU) from a drain beneath the pancreaticojejunostomy for more than 1 week. Pancreatic fistula was defined as prolonged discharge of &gt; 50 ml/day with a high amylase content (&gt; 1000 IU) for more than 1 week.</td>
</tr>
<tr>
<td>Nagakawa and co-workers</td>
<td>Leakage of pancreatic juice occurred when the amylase level in the drainage from the anastomotic site between the pancreas and the intestine was ≥ 1000 IU/l. Suture insufficiency occurred when intestinal contents, particularly bile, were contaminated.</td>
</tr>
<tr>
<td>Reissman and co-workers</td>
<td>Pancreaticojejunostomy anastomotic leak was defined as:</td>
</tr>
<tr>
<td></td>
<td>• recovery of &gt; 40 ml/day of amylase-rich fluid (&gt; 10 times the normal plasma level) from the peripancreatic drains for more than 7 days, or</td>
</tr>
<tr>
<td></td>
<td>• radiologically demonstrable leakage.</td>
</tr>
<tr>
<td>Roder and co-workers</td>
<td>Clinical symptoms (fever, elevated leucocyte count, sepsis), radiographically documented leaks or a fluid collection in the surgical drain adjacent to the pancreaticojejunostomy of &gt; 50 cm³ during the entire course after surgery, with an amylase content exceeding triple the respective serum amylase concentration, was considered a fistula of the pancreaticojejunostomy.</td>
</tr>
</tbody>
</table>
### Definition and Measurement of Anastomotic Leak

**TABLE 15 Overview of the 15 Included Studies of Hepatobiliary Surgery**

<table>
<thead>
<tr>
<th>Study</th>
<th>Investigation</th>
<th>Definition or Cut-off Value Indicative of Leak</th>
<th>Timing*</th>
</tr>
</thead>
<tbody>
<tr>
<td>Ah Chong and co-workers⁹¹⁷†</td>
<td>Retrograde cholangiogram</td>
<td>–</td>
<td>Routine, timing not specified</td>
</tr>
<tr>
<td>Bottger and co-workers²⁴⁵</td>
<td>Drain fluid amylase content</td>
<td>&gt; 2000 units/l amylase</td>
<td>Routine, timing not specified († daily)</td>
</tr>
<tr>
<td>Chou and co-workers²⁴⁶</td>
<td>Drain fluid amylase content</td>
<td>&gt; 50 ml/day drainage fluid for over 2 weeks</td>
<td>Daily</td>
</tr>
<tr>
<td>Davidson and co-workers²⁵⁶</td>
<td>Endoscopic retrograde cholangiogram</td>
<td>Presence of leak on cholangiography</td>
<td>Routine, day 10 to 14</td>
</tr>
<tr>
<td>Evans and co-workers²⁴⁸</td>
<td>Drain fluid amylase content</td>
<td>&gt; 50 ml/day fluid with amylase content &gt; 1000 units/l</td>
<td>Daily</td>
</tr>
<tr>
<td>Gupta and co-workers²⁴⁹</td>
<td>Biliary nucleotide scintigraphy and trans-hepatic cholangiography</td>
<td>Presence of bile from drains</td>
<td>Not specified</td>
</tr>
<tr>
<td>Hamanaka and Suzuki²⁵⁰</td>
<td>Gastrografin swallow</td>
<td>X-ray swallow and clinical signs</td>
<td>Day 7</td>
</tr>
<tr>
<td>Hardy and co-workers²⁴²</td>
<td>Cholangiography or percutaneous transhepatic cholangiography</td>
<td>X-ray investigations</td>
<td>Day 7</td>
</tr>
<tr>
<td>Howard²⁵¹</td>
<td>Serum amylase levels</td>
<td>&gt; 3 times normal serum amylase level</td>
<td>2-Day intervals until day 8 to 12; at least once on day 7 to 10</td>
</tr>
<tr>
<td>Kapoor and co-workers²⁵⁷</td>
<td>Isotope hepatobiliary scanning</td>
<td>X-ray investigation</td>
<td>Timing not specified</td>
</tr>
<tr>
<td>Kayahara and co-workers¹⁹⁷</td>
<td>Drain fluid amylase content</td>
<td>&gt; 3 times the serum amylase level</td>
<td>Timing not specified</td>
</tr>
<tr>
<td>Lowy and co-workers²⁴⁴</td>
<td>Clinical leak: drain fluid amylase content with clinical signs</td>
<td>&gt; 2.5 times normal serum amylase level and if drain output &lt; 200 ml over a 24-hour period</td>
<td>Day 3</td>
</tr>
<tr>
<td>Matsusue and co-workers²⁴³</td>
<td>Drain discharge volume and amylase content</td>
<td>Sepsis: &lt; 50 ml/day drainage with amylase &lt; 1000 units/l for more than 1 week</td>
<td>Daily</td>
</tr>
<tr>
<td>Nagakawa and co-workers²⁵²</td>
<td>Drain fluid amylase content</td>
<td>&gt; 1000 units/l</td>
<td>Timing not specified</td>
</tr>
<tr>
<td>Reissman and co-workers²⁵³</td>
<td>Drain fluid amylase content</td>
<td>&gt; 40 ml/day amylase-rich fluid (&gt; 10 times normal plasma level) for more than 7 days</td>
<td>Daily</td>
</tr>
<tr>
<td>Roder and co-workers²⁵⁴</td>
<td>Radiologically documented leaks; drain discharge volume and amylase content</td>
<td>&gt; 50 ml total postoperatively with amylase &gt; 3 times normal serum level</td>
<td>Daily</td>
</tr>
</tbody>
</table>

* Routine, where tests were performed on all patients
† The study by Ah Chong and co-workers is included in this table as it included details of hepatobiliary investigation, but is counted in the next section
**BOX 24** Definitions of anastomotic leak from studies of lower gastrointestinal surgery

<table>
<thead>
<tr>
<th>Author and co-workers</th>
<th>Definition</th>
</tr>
</thead>
<tbody>
<tr>
<td>Bokey and co-workers</td>
<td>A significant clinical (general) leak was one that necessitated abdominal re-operation. A subclinical (local) leak was defined as one demonstrated by limited Gastrografin enema or one that resulted in an abscess that discharged either spontaneously or following minor surgical drainage.</td>
</tr>
<tr>
<td>Breen and co-workers</td>
<td>Defined a leak as separation of the pouch and anal anastomosis, as detected clinically or at the retrograde water-soluble contrast study before closure of the ileostomy.</td>
</tr>
<tr>
<td>Burke and co-workers</td>
<td>Diagnosed clinically and suspected if there was a deterioration in general condition, abdominal distension, diarrhoea or blood clot passed per anum, or signs of peritonitis.</td>
</tr>
<tr>
<td>Cornwell and co-workers</td>
<td>Defined as presence or absence of suture line disruption, where a colonic suture line existed.</td>
</tr>
<tr>
<td>Deen and Smart</td>
<td>Anastomotic leak indicated by local or generalised peritonism, tachycardia, fever, frank faecal fistula and anastomotic stricture.</td>
</tr>
<tr>
<td>Dehni and co-workers</td>
<td>Evidence of generalised or pelvic infection associated with symptoms such as abdominal pain, fever, leucocytosis or shock.</td>
</tr>
<tr>
<td>Docherty and co-workers</td>
<td>A clinical leak was defined as a dehiscence at the anastomosis confirmed by re-operation, autopsy, the appearance of faecal material from drains, development of a colocutaneous fistula, or the development of any systemic septic associated local peritoneal signs in the postoperative period.</td>
</tr>
<tr>
<td>Fingerhut and co-workers</td>
<td>Overt leakage diagnosed by faecal matter in the drainage discharge, purulent discharge per anum, sinograms, re-operation or post mortem examination.</td>
</tr>
<tr>
<td>Hallbook and co-workers</td>
<td>Symptomatic anastomotic leakage was evident if any of the following was observed: evidence of abscess on a CT scan or ultrasound; discharge of pus either per anum or through a fistula; and necessity of laparotomy or a trans-anal drainage procedure.</td>
</tr>
<tr>
<td>Hansen and co-workers</td>
<td>Clinically suspicious cases of leakage in the presence of an intra-abdominal abscess, postoperative peritonitis, of faecal-stained discharge from the drains.</td>
</tr>
<tr>
<td>Junger and co-workers</td>
<td>Clinical sign of leak reported as faecal secretion into the drain and diffuse peritonitis with intra-abdominal abscess.</td>
</tr>
</tbody>
</table>
| Karanjia and co-workers | Leaks classified as:  
- minor – produced no clinically significant disturbance but was detected by contrast enema  
- major – produced clinically significant effects such as peritonitis or discharge of a pelvic collection. |
| Kelly and co-workers | Anastomotic dehiscence defined by extravasation of contrast at pouchography. |
| Kessler and co-workers | Clinical manifestations: stool fistulas, local abscess, persisting purulent secretion from drainage, peritonitis. |
| Kracht and co-workers | Clinical leaks, assessed by the presence of an intra-abdominal abscess, postoperative peritonitis or by faecal-stained discharge through the incision or drains. |
| Mann and co-workers | Suspected if the patient showed fever, leucocytosis, persistent ileus, bleeding or discharge, or any other sign of intra-abdominal or pelvic abscess. |
| Merad and co-workers | Anastomotic leak diagnosed by the egress of faecal fluid through drains. |
| Miller and co-workers | Clinical leaks were said to occur if pus appeared from the anus, if faeces appeared in the drain, or if pelvic abscess developed. |
| Moore and co-workers | Clinically significant (generalised) anastomotic leak was defined as that necessitating urgent laparotomy. |
| Pakkastie and co-workers | Clinical signs of faecal discharge from the wound or drain, pelvic sepsis, postoperative fever or septicemia. |
| Petersen and co-workers | A clinical leak was defined as an anastomotic dehiscence confirmed by either the appearance of faeces from the wound, the drains, or development of a colocutaneous or rectocutaneous fistula. |
| Redmond and co-workers | Clinical leakage was defined as a breakdown in the double-spaced anastomosis resulting in clinical signs of peritonitis or septic shock such that the patient required laparotomy. |
| Sagar and co-workers | Discharge of faeces from the drain site or the presence of an abscess in close proximity to the anastomosis and localised or generalised peritonitis with tenderness, fever and leucocytosis. |
| Santos and co-workers | Investigation was undertaken into the presence of fever, abdominal or sacral pain, tenesmus, or clinical signs of localised or generalised peritonitis. |
| Slim and co-workers | Clinical anastomotic leak: faecal matter or pus in the drainage discharge, purulent discharge from the anus, intra-abdominal abscess, re-operation for peritonitis. |
in a cohort of 1115 patients based on clinical assessment of leak without radiological confirmation. Results were presented according to individual criteria of stool fistula, local abscess and peritonitis, as well as overall anastomotic leak rate. The presentation of individual features of clinical leakage allows comparison with other centres and institutions.

A second German study group, the Laparoscopic Colorectal Surgery Group, was formed in 1995 with the aim to produce scientifically evaluable data from German-speaking countries in Europe. In 1999 they published results from a large multi-centre observational study of laparoscopic and laparoscopically assisted colorectal anastomoses and the associated risk of postoperative anastomotic leak, based on data from 24 surgical centres in Germany, Austria and Switzerland. Anastomotic leak was one of the most common postoperative complications, with 46 clinical leaks from 949 anastomoses (5%). However, despite this European effort to produce evaluable scientific data, anastomotic leak was not defined and no details are given of how leaks were detected or investigated, either clinically or radiologically.

In a review of 35 publications on the complications of primary repair of colon injury, Curran and Borzotta reported on anastomotic leak rates and optimum management of patients. No details were given on how anastomotic leak was defined, detected or confirmed within the primary literature.

Table 16 lists the clinical features and investigations from the 52 included lower gastrointestinal studies. Fingerhut and co-workers and Miller and co-workers both published two separate studies, but used their own definitions on both occasions. The majority of studies used contrast (water soluble) radiography, either routinely or when leak was suspected. The timing of administration of routine contrast ranged from day 4 to day 14 postoperatively. However, overall assessment of the timing of administration was difficult to interpret as some studies conducted investigations in patients when leak was suspected and then presented results combined with those from routine testing.

One UK-based randomised trial of mechanical bowel preparation altered the policy of routine contrast studies half way through the trial, after two of six anastomotic leaks developed immediately after administration of contrast. These complications led to policy revision, and contrast investigation was subsequently performed only when leak was suspected on clinical grounds.

**Studies of methods and accuracy of detection of anastomotic leak**

Few articles were found that related to the detection of anastomotic leak, and those that were photocopied for appraisal were often found to be single case reports, small case series, retrospective in design or examples of imaging techniques performed after specialised surgical interventions. Some studies focused on prediction of leak rather than detection of anastomotic leak. A short overview of the prediction (Box 25) and detection literature is given below.

**Detection of anastomotic leak**

Barium sulphate is reported to be the optimal contrast material in the detection of anastomotic leaks in the gastrointestinal tract, but extravasation of barium onto surrounding membranes and tissues can lead to chemical irritation and, if within the peritoneal or mediastinal cavity, result in chemical peritonitis or mediastinitis. Barium is a potent adjuvant material for the establishment of peritonitis in conjunction with bacterial contamination. Because of the risk of mediastinitis associated with the use of barium sulphate, water-soluble contrast is generally used to detect upper gastrointestinal leaks, although it is less radiopaque than barium sulphate and thus less effective in detecting small leaks.

Gollub and Bains examined the safety and efficacy of diluted (50%) barium sulphate contrast to detect oesophageal anastomotic leaks in a small group (n = 12) of patients undergoing oesophagograms after oesophagectomy. Leak volumes were calculated as the product of the approximate height, width and depth (in centimetres) on frontal and lateral X-ray projections. The safety and efficacy of barium were determined by two factors:

- retention of barium in the mediastinum that would interfere with subsequent patient care
- the development of mediastinitis, characterised by fever, sepsis, chest pain, a widened mediastinal shadow on chest radiographs, chart review for symptoms, white blood cell count and clinical course.

Few adverse effects resulted from using diluted barium rather than water-soluble contrast material, and the authors recommended that this mixture
### TABLE 16 Overview of 52 studies of lower gastrointestinal surgery

<table>
<thead>
<tr>
<th>Study</th>
<th>Definition</th>
<th>Clinical features</th>
<th>Investigation</th>
<th>Timing</th>
</tr>
</thead>
<tbody>
<tr>
<td>Ah Chong and co-workers²¹⁷</td>
<td>–</td>
<td>–</td>
<td>Water-soluble contrast</td>
<td>When suspected</td>
</tr>
<tr>
<td>Ambrossetti and co-workers²⁸⁷</td>
<td>–</td>
<td>Clinical leaks (undefined)</td>
<td>Water-soluble contrast</td>
<td>Routine, days 9 to 11</td>
</tr>
<tr>
<td>Biondo and co-workers²⁸⁸</td>
<td>–</td>
<td>Clinical leaks (undefined)</td>
<td>Unspecified contrast</td>
<td>When suspected</td>
</tr>
<tr>
<td>Bokey and co-workers²⁶⁰</td>
<td>✔</td>
<td>Gastrografin enema, discharge spontaneously or following drainage</td>
<td>Water-soluble contrast, abdominal re-operation</td>
<td>When suspected</td>
</tr>
<tr>
<td>Bouillot and co-workers²⁸⁹</td>
<td>–</td>
<td>–</td>
<td>Unspecified radiography</td>
<td>Unclear</td>
</tr>
<tr>
<td>Breen and co-workers²⁶¹</td>
<td>✔</td>
<td>Abdominal distension, diarrhoea, blood clot passed per anum, peritonitis</td>
<td>Water-soluble contrast</td>
<td>Routinely prior to stoma closure</td>
</tr>
<tr>
<td>Burke and co-workers²⁶²</td>
<td>✔</td>
<td>Abdominal distension, diarrhoea, blood clot passed per anum, peritonitis</td>
<td>Water-soluble contrast</td>
<td>Routinely on day 7 in the first half of the study, and then changed to when leak suspected</td>
</tr>
<tr>
<td>Cornwell and co-workers²⁶³</td>
<td>✔</td>
<td>Clinical leaks (undefined)</td>
<td>Surgical re-exploration, CT scan or water-soluble contrast</td>
<td>Variable</td>
</tr>
<tr>
<td>Dayton and Larsen²⁹⁰</td>
<td>–</td>
<td>Lower abdominal pain, tachycardia, fever, abdominal distension, peritonitis, pneumoperitoneum, sepsis</td>
<td>–</td>
<td>Unclear</td>
</tr>
<tr>
<td>De Wever and co-workers²⁹¹</td>
<td>–</td>
<td>Pelvic abscess and sepsis</td>
<td>Endoscopy and unspecified radiological test</td>
<td>3 to 4 months</td>
</tr>
<tr>
<td>Debus and co-workers²⁹²</td>
<td>–</td>
<td>–</td>
<td>Barium contrast</td>
<td>When suspected</td>
</tr>
<tr>
<td>Deen and Smart²⁶⁴</td>
<td>✔</td>
<td>Local or generalised peritonism, tachycardia, fever, frank faecal fistula and anastomotic stricture</td>
<td>Unspecified radiography</td>
<td>When suspected</td>
</tr>
<tr>
<td>Dehni and co-workers²⁶⁵</td>
<td>✔</td>
<td>Generalised or pelvic infection associated with abdominal pain, fever, leucocytosis or shock</td>
<td>Water-soluble contrast, imaging or re-operation</td>
<td>Routine contrast study 8 to 10 weeks prior to stoma closure and also when suspected</td>
</tr>
<tr>
<td>Docherty and co-workers²⁶⁶</td>
<td>✔</td>
<td>Faecal material from drains, colocutaneous fistula, systemic sepsis associated with local peritoneal signs</td>
<td>Water-soluble contrast, re-operation</td>
<td>Routine contrast study between day 4 and 14</td>
</tr>
<tr>
<td>Fingerhut and co-workers²⁶⁷,²⁶⁸</td>
<td>✔</td>
<td>Faecal matter in drainage discharge, purulent discharge per anum</td>
<td>Water-soluble contrast, sinograms</td>
<td>Routine contrast study, day 7</td>
</tr>
<tr>
<td>Flohr and co-workers²⁹³</td>
<td>–</td>
<td>–</td>
<td>Multiple investigations performed (ileal neo-bladder); intravenous pyelogram and voiding cystogram at 1 year</td>
<td>Routine intravenous pyelogram and cystogram at 1 year</td>
</tr>
</tbody>
</table>

* These authors published two studies

---

continued
### TABLE 16 contd Overview of 52 studies of lower gastrointestinal surgery

<table>
<thead>
<tr>
<th>Study</th>
<th>Definition</th>
<th>Clinical features</th>
<th>Investigation</th>
<th>Timing</th>
</tr>
</thead>
<tbody>
<tr>
<td>Hallbook and co-workers\textsuperscript{269}</td>
<td>✔ Discharge of pus per anum or through a fistula</td>
<td>Digital and endoscopic examination, contrast, re-operation, CT scan</td>
<td>Routine contrast study prior to stoma closure</td>
<td></td>
</tr>
<tr>
<td>Hansen and co-workers\textsuperscript{270}</td>
<td>✔ Intra-abdominal abscess, postoperative peritonitis, faecal-stained discharge from drains</td>
<td>Unspecified radiography</td>
<td>When suspected</td>
<td></td>
</tr>
<tr>
<td>Hida and co-workers\textsuperscript{294}</td>
<td>–</td>
<td>Water-soluble contrast</td>
<td>Routinely at 2 months</td>
<td></td>
</tr>
<tr>
<td>Hrung and co-workers\textsuperscript{295}</td>
<td>– Abdominal pain, fever, abdominal distension</td>
<td>Water-soluble contrast and/or barium contrast</td>
<td>Routine imaging at 8 to 12 weeks</td>
<td></td>
</tr>
<tr>
<td>Hulten\textsuperscript{296}</td>
<td>– Signs of leakage or sepsis</td>
<td>Endoscopy and radiological examination</td>
<td>Routine contrast 6 to 8 weeks prior to closure of the ileostomy or when suspected</td>
<td></td>
</tr>
<tr>
<td>Iversen and co-workers\textsuperscript{297}</td>
<td>–</td>
<td>Water-soluble contrast</td>
<td>When suspected</td>
<td></td>
</tr>
<tr>
<td>Junger and co-workers\textsuperscript{271}</td>
<td>– Faecal secretion in drain, diffuse peritonitis</td>
<td>Lipopolysaccharide concentration</td>
<td>Lipopolysaccharide assessed daily</td>
<td></td>
</tr>
<tr>
<td>Karanja and co-workers\textsuperscript{272}</td>
<td>✔ Peritonitis, discharge of a pelvic collection</td>
<td>Water-soluble contrast</td>
<td>When suspected</td>
<td></td>
</tr>
<tr>
<td>Kartheuser and co-workers\textsuperscript{298}</td>
<td>–</td>
<td>Unspecified roentgenologic investigation under general anaesthetic</td>
<td>2 months</td>
<td></td>
</tr>
<tr>
<td>Kelly and co-workers\textsuperscript{273}</td>
<td>✔ –</td>
<td>Unspecified contrast</td>
<td>5 to 6 weeks prior to closure of stoma</td>
<td></td>
</tr>
<tr>
<td>Kessler and co-workers\textsuperscript{274}</td>
<td>– Stool fistulas, local abscess, persisting purulent secretion from drainage, peritonitis</td>
<td>Unspecified radiological tests, methylene blue test</td>
<td>When suspected</td>
<td></td>
</tr>
<tr>
<td>Kockerling and co-workers\textsuperscript{285}</td>
<td>– Clinical leaks (undefined)</td>
<td>Unspecified</td>
<td>Unspecified</td>
<td></td>
</tr>
<tr>
<td>Kracht and co-workers\textsuperscript{275}</td>
<td>✔ Intra-abdominal abscess, postoperative peritonitis, faecal-stained discharge through incision or drains</td>
<td>Water-soluble contrast and/or barium contrast</td>
<td>Routine contrast, day 8</td>
<td></td>
</tr>
<tr>
<td>Mann and co-workers\textsuperscript{276}</td>
<td>✔ Fever, leucocytosis, persistent ileus, bleeding or discharge, intra-abdominal or pelvic abscess</td>
<td>Water-soluble contrast</td>
<td>When suspected</td>
<td></td>
</tr>
<tr>
<td>Merad and co-workers\textsuperscript{278}</td>
<td>✔ Faecal fluid through drains</td>
<td>Water-soluble contrast and/or barium contrast</td>
<td>Routine contrast, day 8</td>
<td></td>
</tr>
<tr>
<td>Merad and co-workers\textsuperscript{277}</td>
<td>– Faecal fluid through drains</td>
<td>Water-soluble contrast and/or barium contrast</td>
<td>Routine contrast, day 7</td>
<td></td>
</tr>
<tr>
<td>Miller and co-workers\textsuperscript{279,280}</td>
<td>✔ Pus from anus, faeces in drain, pelvic abscess</td>
<td>Water-soluble contrast</td>
<td>Routine contrast, day 10</td>
<td></td>
</tr>
<tr>
<td>Moore and co-workers\textsuperscript{258}</td>
<td>✔ Clinical leaks (undefined)</td>
<td>Unspecified radiological examination</td>
<td>Routine prior to stoma closure</td>
<td></td>
</tr>
</tbody>
</table>

\textsuperscript{1} These authors published two studies
### TABLE 16 contd  Overview of 52 studies of lower gastrointestinal surgery

<table>
<thead>
<tr>
<th>Study</th>
<th>Definition</th>
<th>Clinical features</th>
<th>Investigation</th>
<th>Timing</th>
</tr>
</thead>
<tbody>
<tr>
<td>Norris and co-workers</td>
<td></td>
<td>Faecal discharge from wound or drain, pelvic sepsis, postoperative fever or septicema</td>
<td>Unspecified imaging or re-operation</td>
<td>When suspected</td>
</tr>
<tr>
<td>Pakkastie and co-workers</td>
<td>✔</td>
<td>Faecal discharge from faecal discharge from wound or drain, pelvic sepsis, postoperative fever or septicema</td>
<td>Water-soluble contrast</td>
<td>Routine contrast, day 7 to 10</td>
</tr>
<tr>
<td>Petersen and co-workers</td>
<td>✔</td>
<td>Faeces from wound or drains, colocutaneous or rectocutaneous fistula</td>
<td>Water-soluble contrast</td>
<td>When suspected</td>
</tr>
<tr>
<td>Redmond and co-workers</td>
<td>✔</td>
<td>Clinical signs of peritonitis or septic shock that required re-operation</td>
<td>Water-soluble contrast</td>
<td>Routine contrast, day 10 to 12</td>
</tr>
<tr>
<td>Richard and co-workers</td>
<td></td>
<td>Faeces from wound or drains, colocutaneous or rectocutaneous fistula</td>
<td>Water-soluble contrast or CT with contrast</td>
<td>Unspecified</td>
</tr>
<tr>
<td>Sagar and co-workers</td>
<td>✔</td>
<td>Discharge of faeces from drain site, abscess in close proximity to the anastomosis, localised or generalised peritonitis with tenderness, fever, leucocytosis</td>
<td>Water-soluble contrast</td>
<td>Routine contrast, day 5 to 7</td>
</tr>
<tr>
<td>Santos and co-workers</td>
<td>✔</td>
<td>Fever, abdominal or sacral pain, tenesmus, signs of localised or generalised peritonitis</td>
<td>Unspecified radiographic examination</td>
<td>When suspected</td>
</tr>
<tr>
<td>Slim and co-workers</td>
<td>✔</td>
<td>Faecal matter or pus in drainage discharge, purulent discharge from anus, intra-abdominal abscess, peritonitis</td>
<td>Water-soluble contrast, re-operation for peritonitis</td>
<td>When suspected</td>
</tr>
<tr>
<td>Stewart and co-workers</td>
<td></td>
<td>Faecal discharge from drain</td>
<td>Unspecified</td>
<td>Unspecified</td>
</tr>
<tr>
<td>Solomon and co-workers</td>
<td></td>
<td></td>
<td>Pouchography, endoluminal transpouch ultrasonography, CT</td>
<td>Unspecified</td>
</tr>
<tr>
<td>Sugerman and Newsome</td>
<td></td>
<td>Faecal drainage</td>
<td>Water-soluble contrast</td>
<td>Unspecified</td>
</tr>
<tr>
<td>Tagart</td>
<td></td>
<td></td>
<td>Limited barium contrast</td>
<td>Routine contrast, day 14</td>
</tr>
<tr>
<td>Thompson and co-workers</td>
<td>✔</td>
<td>Clinical leaks (undefined)</td>
<td>None</td>
<td>Unspecified (not done routinely)</td>
</tr>
<tr>
<td>Watson and co-worker</td>
<td></td>
<td></td>
<td>Water-soluble contrast</td>
<td>When suspected</td>
</tr>
<tr>
<td>Wexner and co-workers</td>
<td></td>
<td></td>
<td>Unspecified contrast</td>
<td>Routine contrast prior to stoma closure, time not specified</td>
</tr>
<tr>
<td>Wheeler and Gilbert</td>
<td></td>
<td></td>
<td>Water-soluble contrast</td>
<td>Routine contrast, day 8</td>
</tr>
</tbody>
</table>
Definition and measurement of anastomotic leak

could be used for routine oesophagography after oesophagectomy. When small clinical leaks are suspected, however, the authors recommend that water-soluble contrast should continue to be used.

One Indian study compared the use of water-soluble contrast with drinking water on a series of 25 consecutive patients after cervical oesophagogastric anastomosis to assess postoperative leakage. All patients underwent a contrast (Gastrografin) study on the day 5 postoperatively, followed by ‘test feeding’ with drinking water in the surgical ward. The drain site dressings were removed and the drain site

### BOX 25 Literature on the prediction of anastomotic leak

Different methods have been used to predict anastomotic leak, including the measurement of haemostatic markers, \(^{225,509}\) the assessment of the blood supply and tissue oxygen tension \((P_{\text{O}_2})\) perioperatively\(^{225,251,309-311}\) and the quantification of various factors in drainage secretions.\(^{271,280}\)

#### Platelet aggregability

Kuwano and co-workers\(^{226}\) prospectively measured the platelet aggregability in 21 patients undergoing oesophageal reconstruction to investigate the relationship between aggregability and occurrence of anastomotic leak. They found a significant difference in the aggregability of platelets on the first postoperative day between patients who developed leakage and those who did not. However, no details were given of how leak was defined or diagnosed. No clinical details and no details about the timing of presentation of patients with anastomotic leak were given.

#### Tissue oxygen tension

Abo and co-workers\(^{310}\) implanted sensors at the anastomotic site in patients undergoing reconstructive surgery of thoracic oesophageal cancer to measure blood and \(P_{\text{O}_2}\) levels, and found that oxygen levels varied significantly in the presence of suture leakage. Schilling and co-workers\(^{251}\) conducted a similar uncontrolled, observation study to measure gastric perfusion before, during and after gastric tube formation in patients undergoing thoracic oesophagectomy. Junger and co-workers\(^{271}\) evaluated the total daily secreted lipopolysaccharide concentration in drainage fluid to assess whether it could be used as a predictor of anastomotic leak. The lipopolysaccharide concentrations on the first two postoperative days were found to vary considerably, with 20–25\% false-positive results and 50\% false-negative results when a lipopolysaccharide threshold of 1500 pg/ml was used. After the third postoperative day, all patients with anastomotic leak had lipopolysaccharide levels exceeding 3000 pg/ml; the corresponding levels in patients without anastomotic leak were consistently lower than 2000 pg/ml. The authors suggested that a cut-off value could be used to diagnose anastomotic leak, where values of > 250 ng/day or > 5000 pg/ml lipopolysaccharide indicate leak.

#### Drain fluid assessment

Miller and co-workers\(^{280}\) measured lysozyme activity in 42 patients undergoing low anterior resection of the rectum. Drain fluid secretions were collected daily from day 1 to day 4 postoperatively to determine lysozyme concentrations. A water-soluble contrast enema was undertaken on day 10 postoperatively and any leak of contrast material was interpreted as a ‘radiological leak’. Clinical leaks were said to occur if pus appeared from the anus, if faeces appeared in the drain or if pelvic abscess developed. The mean lysozyme activity was increased in those patients with clinically (18 mg/dl on day 1 postoperatively) and radiologically (15.3 mg/dl on day 1 postoperatively) detected dehiscence.

Simmen and co-workers\(^{437}\) analysed pH, \(P_{\text{O}_2}\) and \(P_{\text{CO}_2}\) levels in peritoneal fluid obtained from drains placed during laparotomy in 55 patients undergoing emergency surgery. Seven patients developed anastomotic leak or abscess formation, which was assessed clinically (fever, bowel function) and biochemically (leucocytes and C-reactive protein). Anastomotic integrity was also assessed by water-soluble contrast in patients suspected of having leakage. Significant differences were found in the values of pH, \(P_{\text{O}_2}\) and \(P_{\text{CO}_2}\) between patients who developed anastomotic complications and those who did not, and this was found before clinical symptoms were evident. Sensitivity values were less than 50\%, but specificity was over 94\% for each of the three parameters and increased to 100\% when pH and \(P_{\text{O}_2}\) values were combined.

Yamaguchi and co-workers\(^{438}\) used litmus paper to detect leaking pancreatic ductules in a small series of 10 patients undergoing pancreatocystostomy. Red litmus paper was used to detect transected pancreatic ductules on the cut surface of the pancreatic stump. After several minutes the paper showed diffuse blue reactants that corresponded to transected pancreatic ductules, which could then be transfixed.

**Note:** Although the majority of these studies reported significant findings, many were small, observational, uncontrolled studies. For example, the study by Yamaguchi and co-workers on litmus paper reported that no major postoperative leakages occurred, but the sample size of ten patients does not allow firm conclusions to be drawn about the predictive benefits of the use of litmus paper intraoperatively.
in the neck was closely observed by the surgical team whilst the patient took sips of water, and results were compared to those from contrast radiography. The results are given in Table 17. One patient was unable to tolerate contrast. One of three leaks identified by contrast study were clinically apparent on test feeding. The authors concluded that routine use of contrast may not be necessary as the number of leaks detected radiologically is often in excess of the number presenting clinically, and may be false-positive results or too small to be a problem clinically.

One UK study examined the use of intraoperative controlled water testing of anastomoses in 102 patients undergoing left-sided colorectal resection. Saline was used to test the integrity of each anastomosis, via a manometer, with a maximum distending pressure of 30 cmH₂O exerted. A water-tight anastomosis was achieved in 79% of patients on one leakage test, and this rose to 95% on the second test after further suturing. Although this method highlighted patients with an imperfect anastomosis during surgery it did not prevent leak, and although the testing method was helpful it did not guarantee an intact anastomosis for the remainder of the postoperative period.

Bischof and co-workers examined the role and value of CT cystography in the diagnosis of duodenal leaks from kidney–pancreas transplants in a small group of 18 patients. This surgery entailed formation of an anastomosis between a transplanted duodenal segment to the dome of the bladder, and pancreatic juice is subsequently excreted into the urine. A leak from the duodenal segment is difficult to diagnose but is a serious and significant complication. They defined leak, using CT cystography, if extravescical contrast material or air was demonstrated, or if the amount of pelvic or peripancreatic fluid on postvoiding scans significantly increased. A leak was considered to be an early complication if it occurred within the first 5 weeks (35 days) and a late complication if it occurred after that time. The CT protocol comprised plain CT scans, CT cystograms with the bladder fully distended by iodinated contrast material and air and, if found to be negative, CT scan after voiding. The CT scans were compared to the gold standard of findings at surgery, cystoscopy and multiple clinical follow-up examinations. Overall, the diagnosis based on CT cystography was correct in 23 of 24 studies, with 11 true-positive, 12 true-negative, one false-negative and no false-positive results. The overall values for CT cystography were sensitivity 92%, specificity 100% and accuracy 96%. This was the only study included that formally evaluated the accuracy of a test for the diagnosis of anastomotic leak.

Critical assessment of the properties of anastomotic leak

Validity and reliability of the definition of anastomotic leak

There is no single definition of anastomotic leak. This review identified 46 definitions from 107 studies. Less than half of the studies clearly defined leakage, with only nine definitions reported in 52 upper gastrointestinal studies (23%), 13 definitions reported in 15 hepatopancreaticobiliary studies (87%) and 27 definitions reported in 52 studies of lower gastrointestinal surgery (52%).

At a workshop convened in 1991, the UK SISG proposed a number of postoperative definitions for use in clinical audit and to form the basis for meaningful comparisons. Anastomotic leak was defined as: “the leak of luminal contents from a surgical join between two hollow viscera. The luminal contents may emerge either through the wound or at the drain site, or they may collect near the anastomosis, causing fever, abscess, sepsicaemia, metabolic disturbance and/or multiple-organ failure. The escape of luminal contents from the site of the anastomosis into an adjacent localised area, detected by imaging, in the absence of clinical symptoms and signs should be recorded as a subclinical leak”.

Definitions were also given for surgical wound infection, sepsicaemia, generalised peritonitis, chest infection, urine infection and pelvic abscess. The SISG definition for anastomotic leak is generic, in that it applies to any anastomosis of the gastrointestinal tract. Unlike the SISG definition of surgical wound infection, which

---

**TABLE 17 Results from the study by Goel and co-workers**

<table>
<thead>
<tr>
<th>Water-soluble contrast study</th>
<th>Test feeding</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Leak</td>
<td>No leak</td>
</tr>
<tr>
<td>Leak</td>
<td>1</td>
<td>2</td>
</tr>
<tr>
<td>No leak</td>
<td>1</td>
<td>20</td>
</tr>
<tr>
<td>Unsatisfactory</td>
<td>1</td>
<td>0</td>
</tr>
<tr>
<td>Total</td>
<td>3</td>
<td>22</td>
</tr>
</tbody>
</table>
was implemented in several UK studies and referred to by others, the SISG definition of anastomotic leak was not cited or referred to in any of the 240 appraised articles of surgical anastomosis. Yet this definition is comprehensive in that it outlines what a leak comprises (“the leak of luminal contents from a surgical join between two hollow viscera”), contains common signs and symptoms indicative of leakage (“causing fever, abscess, septicaemia, metabolic disturbance and/or multiple organ failure”) and also includes a definition for subclinical leak (“detected by imaging in the absence of clinical symptoms and signs”). However, it could be modified to include parameters for fever and metabolic disturbance. Postoperative septicaemia was defined by the SISG as: “when rigors occur together with one or more of the following signs: fever, higher than 38°C on more than one occasion in 24 hours, and/or hypotension, and/or oliguria. Objective evidence of the source of infection and laboratory confirmation of viable microorganisms (bacteriaemia) or their products (exo- or endotoxins, antigen or antibody) in the blood are desirable to confirm the diagnosis”.

Recent literature suggests that the term ‘septicaemia’ should be discarded and replaced by ‘systemic inflammatory response syndrome’ (SIRS) or ‘sepsis’, where SIRS is due to infection. A definition of anastomotic leak that includes sepsis should, therefore, use these parameters. In 1991, SIRS was defined by the American College of Chest Physicians/American Society of Critical Care Medicine as two or more of the following:

- temperature > 38°C or < 36°C
- tachycardia > 90 beats/minute
- respiratory rate > 20 beats/minute or PaCO₂ < 4.3 kPa
- white cell count > 12 × 10⁹/l or < 4 × 10⁹/l or < 10% immature band forms.

**Clinical assessment of oesophagogastric anastomotic leak**

In 1994, Bardini and co-workers highlighted the variation (0–35%) in the incidence of anastomotic leak following oesophageal surgery, suggesting this may in part be due to lack of an agreement on the definition of oesophageal anastomotic leak. Of the nine definitions of upper gastrointestinal leakage, six specified clinical features: evidence of haematoma or seroma formation at the neck wound; leaks with perianastomotic collection; saliva in the chest drain, mediastinitis or abscess, pneumothorax, empyema; septicaemia; retarded bowel function, fever, localised abdominal fluid; and peritonitis. Clearly, there is no consensus of opinion on the clinical presentation of an upper gastrointestinal anastomotic leak.

Three classification systems were identified, in the studies by Csendes and co-workers (type I, type II), Isozaki and co-workers (minor leak, major leak) and Nambirajan and co-workers (incidental leak, minor leak, major leak). Both levels of the system proposed by Csendes and co-workers are based on findings from radiological studies, with no consideration or inclusion of clinical signs or symptoms. Isozaki and co-workers classified severity as minor and major, but this grading was almost identical to that proposed by Csendes and co-workers, except for the inclusion of “or cases in whom symptoms of peritonitis required the insertion of a new drain” under major leakage. The Csendes classification system is only of value for grading radiologically detected leaks.

In contrast, Nambirajan and co-workers.classed radiological leaks where the patient had no clinical symptoms as incidental, and minor leakage as “saliva in the chest drain but clinically well”. The definition for major leak includes both clinical signs and radiological findings, although it is unclear whether these factors each stand alone or should be used in combination. It is unlikely that saliva in the chest drain is the only possible clinical feature, and therefore the definition may not be valid as it does not include all possible properties of the event (content validity). This definition, however, clearly gives three separate levels of severity of leakage, distinguishing between radiological, clinically insignificant and clinically significant leaks.

**Clinical assessment of hepatopancreaticobiliary leak**

In a review of complications after pancreatectomy, Cunningham and co-workers stated that the most common presentation of a patient with a pancreatic leak is “fever associated with abdominal pain, leucocytosis and often an ileus”, along with increased amylase-rich output from drains. Of the 13 definitions in the hepatopancreaticobiliary literature, only three studies included clinical signs. Hamanaka and Suzuki included “peritonitis, pyrexia and sepsis”, and Roder and co-workers accepted “fever, elevated leucocyte count and sepsis”. Neither of these reports gave upper temperature values for pyrexia or values for leucocyte cell counts. Lowy and co-workers clearly outlined the criteria for clinical leak,
with parameters given for fever (> 38°C), cell count (white blood cell count > 10,000/l) and sepsis (haemodynamic instability requiring transfer to the intensive care unit). This was the only definition of sepsis identified in the literature review, other than that given by the SISG. There appears to be more consensus over the clinical presentation of leaks from hepatopancreatobiliary surgery, in that the common features are pyrexia, an elevated white cell count and sepsis.

In general, the definitions of hepatopancreatobiliary leakage were more precise as they also included quantitative values for volume of drainage and enzyme concentration, although these values varied from paper to paper. The definitions that specify a time after which leakage becomes significant suggest that leakage is accepted as routine but clinically significant before that arbitrary date. In 1998, Cunningham and co-workers states that leakage after pancreatic surgery is usually defined as “amylase rich output from drains that is three to five times the concentration of serum amylase”. Our review identified significant variation in definitional (or diagnostic) levels of amylase concentrations, with one study accepting up to 10 times the normal plasma level for more than a week. Other studies were less specific: Chou and co-workers classed leaks as “amylase rich fluid”. Unlike the upper gastrointestinal literature, no grading or severity scales were identified, although Matsusue and co-workers distinguished between peripancreatic sepsis and pancreatic fistula, and Lowy and co-workers distinguished between a clinical leak and a biochemical leak. Although no formal critical assessment of the definition was conducted, the definition by Lowy and co-workers appeared to be valid (content validity) in that it contained the relevant properties for assessment of leakage, and is likely to be reliable, because the clearly defined parameters reduce the likelihood of intra- and inter-observer variation.

Clinical assessment of lower gastrointestinal anastomotic leak

More clinical features were contained in the definitions of lower gastrointestinal anastomotic leak than in the definitions of upper and hepatopancreatobiliary surgery. The clinical signs and symptoms most frequently cited in the 24 definitions included: signs of localised or generalised peritonitis (n = 12); faecal discharge from the wound and/or drain (n = 11); abscess (n = 10); purulent discharge from the drain, wound or anus (n = 7); and fever (n = 6). None of the studies defined what was meant by ‘peritonitis’. The SISG defined postoperative peritonitis as: “generalised peritonitis is a diffuse inflammation of the peritoneum caused by infective agents or by toxic substances associated with the clinical manifestations of abdominal pain, tenderness and guarding, and subsequently by impaired alimentary tract function. The latter may be absent under certain postoperative conditions, e.g. artificial ventilation, and it is accepted that clinical signs and symptoms may be difficult to interpret in the immediate postoperative period. It is desirable that the diagnosis of peritonitis due to infection is supported by positive bacterial culture of the peritoneal exudate”.

Faecal or purulent discharge from the wound and/or drain is a more definitive measure, as is fever, although none of the six definitions that used fever included a temperature value. Karanjia and co-workers classed leaks as minor or major, whereby minor leaks were detected radiologically without any clinical disturbance, and major leaks led to clinically significant results. Bokey and co-workers differentiated between significant leaks and subclinical leaks. No other severity systems were used, although many authors distinguished between clinical and radiological leaks.

The study by the French Association for Surgical Research conducted routine contrast studies in order to obtain an accurate evaluation of all anastomoses. This allowed reproducible assessment across all 24 surgical centres. Although interpretation of images was conducted by different assessors, this method is more reliable than clinical assessment of signs and symptoms that vary across the centres by multiple surgical teams. Two independent assessors were used to review radiographic investigations in a number of studies, but none measured the extent of agreement or variation between assessors (reproducibility).

There is considerable variation in the terminology related to clinical leakage, including significant clinical (general) leak, overt leakage symptomatic anastomotic leakage and clinically significant generalised anastomotic leakage. Some studies defined a clinical anastomotic leak as that requiring re-operation, while others accepted signs of leak without further surgery. This was also true of the upper gastrointestinal and hepatopancreatobiliary literature, where some authors accepted clinical features (e.g. haematoma) as evidence of leakage but...
Definition and measurement of anastomotic leak

others only accepted a definition of leakage if confirmed on re-operation.207 The term ‘clinical leak’ is not directly comparable across studies. The term ‘re-operation’, which was included in many definitions, is an outcome rather than a definition in itself. Re-operation is arguably an easy outcome to measure, but the timing needs to be clarified, and whether or not it includes surgery after the primary hospitalisation (e.g. patients having had short-stay surgery re-admitted for further surgery with suspected anastomotic leak).

Conclusions

In conclusion, no single validated definition was found of anastomotic leak in the literature. A definition was proposed at a UK consensus workshop in 1991, but no evidence was found of this definition other than the workshop publication. The present review identified 46 definitions from 107 studies. There appears to be little consensus regarding the clinical features of leakage after upper gastrointestinal surgery, and most studies relied on routine contrast investigation in the first postoperative week to detect anastomotic leak. Within hepatopancreaticobiliary surgery, the most common clinical features of anastomotic leakage include fever, elevated leucocyte count, sepsis and an excess of amylase-rich drainage fluid, although values for drainage volume and enzyme concentration varied from study to study. The definitions of anastomotic leak after bowel surgery include the presence of peritonitis (localised or generalised), faecal or purulent drainage from the wound and/or drain, presence of an abscess, and fever. The majority of studies performed routine contrast studies at the end of the first postoperative week to determine anastomotic integrity.

There are no formal evaluations of any of the identified definitions, in terms of content, criterion or construct validity, whereby content, in terms of individual components, has been assessed, or the application of two definitions of anastomotic leak have been applied to the same population at the same time (criterion validity). The lack of a single definition or a gold standard that has been accepted and implemented throughout the surgical field against which to make comparisons hampers any such formal assessment. The majority of the definitions require subjective assessment, particularly the definitions of lower bowel surgery, which tend to comprise clinical signs and symptoms. Although no formal evaluations or quantitative assessments of reliability have been conducted, these definitions will be subject to variation in interpretation, as clinical assessment is likely to be made by more than one member of a surgical team over a period of hospital stay. Assessment and diagnosis of sepsis may differ between junior and senior surgeons. There is significant variation in the definition and interpretation of clinical leakage and, although a single definition is used throughout a study, it is unlikely that comparisons with other centres or institutions will be valid.

There is thus a clear need for surgeons to accept a single, standard definition and grading system that has been demonstrated on scientific evaluation to be valid and reliable, and that distinguishes between radiological, clinical and clinically significant anastomotic leak after gastrointestinal surgery.
DVT is a frequent and serious postoperative complication and is the most common preventable cause of death in hospitalised patients. Early diagnosis of DVT is essential to reduce the risk of pulmonary embolism and post-thrombotic syndrome. This syndrome can include chronic deep vein insufficiency, chronic pain, venous stasis, recurrent cellulitis and lower extremity ulceration. Thromboembolic disease can occur after surgery, trauma or cancer, and can also develop without any of these conditions. The incidence of DVT following orthopaedic surgery is particularly high, with a 20% risk of developing DVT after hip fracture surgery, even when effective prophylactic regimens are implemented. The clinical course of DVT may be complicated by recurrent episodes, by pulmonary embolism and by the development of serious post-thrombotic sequelae, including venous ulceration, debilitating pain and intractable oedema. More than 90% of pulmonary emboli are caused by DVT and most patients who die from pulmonary emboli do so within 30 minutes of the acute event. It has been estimated that 85–95% of pulmonary emboli originate as DVT in the pelvis and lower extremities. As prophylactic measures provide incomplete protection against thromboembolism, detection of clinical and subclinical DVT is an important management strategy because impaired venous function may occur for years after the episode.

The literature on the accuracy of measurement of DVT was markedly different from that of surgical wound infection and anastomotic leak. Diagnosis of surgical wound infection and anastomotic leak are predominantly based on subjective clinical assessment, whereas there is a plethora of diagnostic tests for assessing the presence or absence of thromboembolic disease. The volume of literature related to the definition and measurement of DVT was much larger than for the other adverse events and a greater proportion of retrieved studies fulfilled the inclusion criteria. The search of bibliographic databases retrieved 1239 abstracts, 250 of which were selected for full critical appraisal. Of these, 30 articles were comprehensive literature reviews or meta-analyses on diagnostic accuracy. It was anticipated that more studies would be identified after reference checking and hand-searching. Many different methods of diagnostic assessment were identified from critical appraisal of a small proportion of the total number of studies (n = 64) on the diagnostic accuracy of DVT. It was estimated that it would take at least 3 months to review and appraise this literature on so many different methodologies and that time was not available within the project. It was therefore agreed with the NHS HTA Programme that DVT as an example of a surgical complication might be postponed to a separate systematic review.

This chapter therefore gives only a brief overview of the different diagnostic tests available for the measurement of thromboembolic disease and highlights the important factors to be considered given that future work is recommended in this field.

Diagnostic techniques
At least five separate techniques are available for the detection and diagnosis of DVT. These are, broadly, venography, sonography, plethysmography, radiolabelled fibrinogen scanning and magnetic resonance imaging. Each has advantages and disadvantages, including different degrees of sensitivity and specificity, invasiveness, exposure to ionising radiation, operator dependence and cost. Venography has long been considered the reference standard or gold standard diagnostic test, particularly for lower extremity DVT, but the last decade has seen an increase in the use of sonography and other non-invasive techniques. This is partly due to the perceived invasiveness of venography and improvement in ultrasonographic techniques, with the advent of colour flow and Doppler imaging. Other non-invasive imaging techniques are also currently being developed and tested (e.g. high-resolution infrared thermography and magnetic resonance using gradient-recalled acquisition in steady state (GRASS)). Short descriptions of the main diagnostic tests are given below with a brief note of the types of studies found in the literature.

Clinical examination
Clinical examination of postoperative patients is generally considered insensitive and non-specific for the diagnosis of DVT. The symptoms or clinical
predictors of DVT include the presence of pain, oedema, asymmetry of circumference in the lower limbs and fever. None of the signs and symptoms of DVT are unique and many other disorders can mimic DVT, including superficial thromboembolitis, trauma, post-thrombotic syndrome, oedema and external venous compression due to malignancy. Studies have examined the predictive power of clinical symptoms and the ability of subspecialists to diagnose DVT by clinical examination. Attempts have also been made to devise critical pathways, and regression models have been designed to incorporate clinical signs and symptoms.

**Venography**

Venography was first introduced in 1923 and is the injection of contrast agent, either ionic or non-ionic, into the foot while the superficial veins are occluded by a tourniquet. This allows visualisation of the lower venous system. The main role of ascending phlebography is the diagnosis of acute DVT and chronic venous disease. The advantages of venography include the direct visualisation of the thrombus, which is viewed as a filling defect. The sensitivity of venography is such that it is regarded as the ‘gold standard’ investigation. The disadvantages of venography are that it is invasive, costly and is reportedly unreliable in about 5–15% of cases. Incomplete venous filling and other technical problems lead to inadequate results in up to 5% of cases. Furthermore, venography is also user-dependent, leading to disagreement over the presence of a thrombus in about 10% of cases. Venography can cause discomfort and may occasionally cause DVT. Many approaches to venography have been described, including different patient positioning, different contrast media and the use of tourniquets. Variation in diagnostic methodology may be responsible for differences in DVT rates between clinical centres.

**Ultrasoundography (compression ultrasonography, Doppler, duplex scanning)**

Two-dimensional ultrasonography produces an image of the deep veins, and thrombi can be detected by direct visualisation or vein non-compressibility. The aims of sonography are the determination of the presence of a thrombus, and the evaluation of the extent of the thrombus, its age and attachment to the venous wall. Particular features of interest include whether or not the thrombus is partially or totally occlusive, attached or free-floating. A standard lower extremity compression ultrasound examination includes the common femoral, superficial femoral, popliteal and calf veins.

Compression sonography is performed by applying minimal pressure with the ultrasound probe to the underlying vein at 1–2 cm intervals; veins free of thrombus will collapse, obliterating the entire venous lumen. Some literature suggests that ultrasonography is regarded as the investigation of choice for suspected lower limb DVT.

In 1997, Dauzat and co-workers described the major and minor criteria for diagnosis of DVT by sonography. Major criteria include: the presence of echogenic intraluminal material (reliability depends on contrast resolution and the correct setting of the dynamic range of the B-mode system); vein incompressibility (a vein cannot be collapsed by moderate pressure with the ultrasound probe due to thrombus); and the absence of spontaneous or elicited blood flow (which can be detected with continuous-wave, pulsed (duplex) or colour Doppler ultrasound in proximal veins). Minor criteria include: an enlarged vein diameter; venous wall and valve immobility; upstream and downstream spontaneous blood flow echogenicity; and enlarged collateral veins with increased flow. There is variation on which criteria are used, but vein incompressibility is most frequently used as a diagnosis of DVT.

Limitations of ultrasound include the difficulty in assessing venous thrombi above the inguinal ligament due to pelvic bones and bowel gas, and visualisation is also difficult in the distal popliteal vein. The accuracy of the role of ultrasound in the assessment of calf thrombosis and its use in screening of asymptomatic patients is debatable.

Continuous-wave Doppler sonography was first used in the late 1960s and is the use of Doppler shifts to detect venous blood flow. Doppler examination can be performed using a handheld unit and listening over areas of major venous flow, and has been recommended for use when other objective tests are not available. Doppler examination is reportedly time-consuming, operator dependent and requires access to expensive equipment, but the equipment is portable and easy to use. Simons and co-workers state that colour Doppler improves sensitivity because:

- it provides a two-dimensional image of the pulse-wave Doppler velocity measurement, enabling the reader to visualise responses to augmentation manoeuvres
- it allows the technician to readily distinguish arteries from veins
- it diagnoses non-occlusive thrombus when flow is detected around an anechoic mass within the vein
- it enhances the technician’s ability to discern deep calf veins, which are often small and difficult to distinguish from other structures when examined using the grey-scale technique.

Duplex scanning is the combination of Doppler measurement with the two-dimensional ultrasound image. Katz and co-workers have described duplex ultrasonography as the “simultaneous use of real-time ultrasound imaging combined with pulsed gated Doppler technology where real-time imagining allows direct visualisation of the vein and Doppler instrumentation produces an audible and/or graphic representation of blood flow”. Duplex imaging is reportedly accurate for the diagnosis and surveillance of both above and below knee thrombi. Duplex scanning is safe, painless and can be used in pregnancy. The criteria considered for diagnosis of DVT by duplex imaging include: visualisation of thrombus; absence of spontaneous flow by Doppler ultrasonography; absence of phasicity of flow with respiration; and incompressibility of the vein with probe pressure.

**Plethysmography**

Plethysmography is the detection of impaired venous emptying of the leg. Different techniques include impedance plethysmography, pneumoplethysmography, quantitative air plethysmography, photoplethysmography and light reflection rheography. Impedance plethysmography is the measurement of the rate of venous return when a pneumatic thigh cuff is inflated to obstruct venous flow. The resulting increase in volume of the leg can be monitored as a decrease in the impedance between pairs of electrodes that are applied around the calf. Impedance plethysmography was hailed as an important development due to its high sensitivity and specificity in diagnosing symptomatic proximal thrombi. Impedance plethysmography was widely used in the 1980s, but was gradually replaced in the 1990s by the use of ultrasound. It is non-invasive and portable, and therefore can be used at the patient’s bedside, and is reportedly less sensitive than compression ultrasonography and venography in symptomatic patients. However, the value of the use of impedance plethysmography in asymptomatic patients is debatable.

Light reflection rheography is a non-invasive procedure and is based on the principle of photoplethysmography, whereby near-infrared light is beamed 1–2 mm into the skin from three light-emitting diodes. The volume of blood in the dermal venous plexus determines the absorption and reflection of light via a central detector. Sproule and co-workers have reported that light reflection rheography is a quick and effective screening method for DVT and can reliably exclude DVT.

**d-Dimer assay**

Other non-invasive techniques include the measurement of the plasma level of d-dimer. A large number of assays are now commercially available and, although they each vary in assay principles, all are basically immunochemical methods employing antibodies directed against cross-linked peptides derived from protease (plasmin) digestion of fibrin. There has been a need for a rapid, sensitive and accurate test to exclude DVT in order to prevent patients from having to undergo elaborate, invasive and expensive tests. d-Dimer assays, therefore, are the product of the interaction between plasmin and cross-linked fibrin, where fragments of multiple molecular weights produce differences in immune reactivity. Elevated levels of cross-linked fibrin derivatives are found in DVT, pulmonary embolism, disseminated intravascular coagulation and malignancy. One of the main uses of the d-dimer assay is to exclude venous thromboembolism in patients suspected to have this condition.

There are three broad groups of d-dimer assays: quantitative, semi-quantitative and qualitative. Quantitative assays include d-dimer enzyme-linked immunosorbent assay (ELISA), Asserachrom d-dimer ELISA, TintElize, SimpliRED d-dimer and Vidas d-dimer enzyme-linked fluorescent immunoassay (ELFA). Semi-quantitative assays include Minute Latex, d-dimer Latex and NycoCard d-dimer. Qualitative assays include instant immunoassay (IA). Sensitivity and specificity will vary according to method and upper cut-off value used.

**Magnetic resonance imaging**

The use of magnetic resonance imaging is a relatively recent approach to the diagnosis of DVT. There are reports that magnetic resonance imaging is significantly more accurate than sonography in detecting lower extremity DVT, and it is thought to be as sensitive and specific as contrast venography in the detection of DVT in the thigh. It is reported to excel in the diagnosis
of pelvic thrombi. Although magnetic resonance imaging is an accurate technique and is not operator dependent, it is expensive, availability is often limited and occasionally it is contraindicated.

**Other techniques**

Other diagnostic methods identified in the review included the use of spiral CT, studies of concentration of serum C reactive protein and other techniques using labelled platelets and antibodies.

**Factors for critical appraisal**

Details extracted from individual studies included: study details; aim; population (patients/limbs); gold-standard test; comparative test; patient population; specific anatomical region; values for validity, reliability and accuracy by anatomical region and patient group; and practicality and feasibility of test. Although the critical appraisal was not completed, the review panel identified several important variables during the process of selection and extraction of data from the DVT literature. We recommend consideration of these factors for any future systematic review.

They are:

- the recording of anatomic region (lower limb/upper limb/pelvis)
- patient presentation (symptomatic/asymptomatic)
- completion and success of test (completed/inconclusive/technically inadequate/negative)
- length of follow-up
- whether or not serial screening was conducted
- laboratory cut-off values.

The reasons for the inclusion of the above factors are set out below.

Thromboembolism in the proximal veins carries a higher level of risk of propagation than in the distal veins. Nevertheless, calf vein DVTs also propagate and are a significant cause of pulmonary emboli. The accuracy of each imaging test depends on whether the DVT is situated in the calf, thigh or pelvis. It is not sufficient to extract overall results on the accuracy of a diagnostic test, each article must have results extracted according to region. This factor must be taken into account for overall interpretation and final presentation and we recommend that accuracy is stratified by anatomic region. Similarly, differences exist in the accuracy of individual tests according to risk status. Accuracy of screening of asymptomatic patients, such as the postoperative population, is affected by the level of risk according to the type of procedure (e.g. orthopaedic surgery) and other factors, including previous history and co-morbidity. The accuracy of detection of DVT in symptomatic patients varies according to whether the population under review has symptoms of acute DVT or ongoing symptoms from chronic thrombophlebitic disease.

Other issues related to the reporting of results include whether or not individual studies report success or grade test results (e.g. specify whether thrombosis was detected, inconclusive results, test technically inadequate or result was negative). Often, a further spectrum of investigations is undertaken and the comparison of one test to a gold standard is thus complicated by the introduction of a new diagnostic test and subsequent variables. Finally, results from individual studies of non-invasive D-dimer assays should be recorded with a cut-off value for fibrinogen equivalent units. Such cut-off values vary between laboratories and results cannot be amalgamated unless comparing like with like.

In summary, there are a number of different factors to be considered in the assessment of the diagnosis and measurement of DVT. We also recommend that, where possible, future work in this field should consider the cost of individual investigative tests.
Chapter 8

Surgical mortality monitoring systems

This chapter deals with the monitoring of surgical mortality within surgical information or monitoring systems. In devising a search strategy for monitoring systems about which there might be some critical literature, a preliminary search suggested that there were two main types. First, event-specific monitoring systems, such as those for surgical nosocomial infections and for surgical mortality; and, secondly, national or state routine systems, from which information on surgical patients and postoperative events could be extracted. Articles were included if they described systematic and continuing data collection to monitor postoperative deaths. The vast majority of the literature concerned cardiothoracic surgical mortality, with a smaller but important component on surgical deaths across all specialties. The latter came from continuous audits, confidential enquiries and national or organisational monitoring systems that had a range of purposes. A search for monitoring systems in their own right yielded a mass of articles that did not describe systematic continuous monitoring but offered some useful insights into ways of improving the accuracy and completeness of data collection. The four national routine NHS systems were examined for their potential to act as monitors of surgical mortality.

The criteria we anticipated for a good monitoring system were described in chapter 2. Briefly, these were:

- the use of standard definitions and duration of follow-up (tables of definitions and follow-up are given in chapter 9)
- a denominator from which to calculate rates
- inclusion of risk factors
- data collection and processing that gives complete, accurate and reliable ascertainment of events and meets the requirements of data protection legislation
- output and feedback that is timely and user-friendly.

These criteria were applied to the literature described in this chapter. This chapter describes examples of monitoring systems of mortality within general, cardiothoracic and vascular surgery, and gives an overview of generic, routine data collection systems that can be used to monitor surgical adverse events, in particular surgical mortality, both within the UK and in other countries.

Surgical mortality monitoring systems within general surgery

Two main examples of surgical mortality monitoring systems exist in the UK, the NCEPOD138 and the SASM.14

The National Confidential Enquiry into Peri-Operative Deaths

The NCEPOD138 is the national successor to the Confidential Enquiry into Peri-Operative Deaths (CEPOD), which was set up in 1986 by the Association of Surgeons and the Association of Anaesthetists and extended to the whole of the UK in 1989, except Scotland, which is covered by the SASM.14 The NCEPOD is an independent limited company and registered charity and is funded in part by: the Department of Health of England through the National Institute for Clinical Excellence; the Departments of Health for Wales, Northern Ireland, the Isle of Man, Jersey and Guernsey; and contributions from private sector hospitals. Hospitals from all the above, together with the Defence Medical Services Hospitals, provide the participants to the Enquiry. In 1999, the NCEPOD reported a participation rate by surgeons and anaesthetists of approximately 78%.348 The aim of the NCEPOD is to produce an analysis of common and potentially remediable factors in a group of surgical patients who die, not to review individual deaths and assign a cause; thus its focus is primarily epidemiological rather than clinical.

The NCEPOD collects basic details on all deaths occurring in hospital within 30 days of a surgical procedure. It covers all hospital episodes, but not deaths postdischarge. A surgical procedure is defined by the NCEPOD as “any procedure carried out by a surgeon or gynaecologist, with or without an anaesthetist, involving local, regional or general anaesthesia or sedation”. The NCEPOD does not include cardiac deaths or maternal deaths, which are covered by the...
National Confidential Enquiry into Maternal Deaths. There is no denominator population, in that the Hospital Episode Statistics for England do not record the number of operations undertaken. In the first NCEPOD report an estimate was made of the expected number of surgery-related deaths as an attempt to gauge completeness of reporting, but it is not possible to calculate operation-specific case-fatality rates.

Data on each identified death (hospital or NHS number, date of birth, sex, date of operation, date of death, consultant in charge, anaesthetist) are submitted to the NCEPOD by a designated local reporter within each hospital. The NCEPOD is registered with the Data Protection Commissioner and abides by the Data Protection Principles. All paper records are shredded once a report has been published and only anonymous, aggregated data are retained on computer. In 1998, the return rate approached 80%. Each year a sample of the total deaths, complete and accurate data were obtained. In this hospital, coding occurred immediately after the death rather than after the discharge letters had been written. A trial of immediate coding in the first hospital improved results for 3 months and then ascertainment gradually dropped again. Cook and co-workers calculated the sensitivity and specificity of the ascertainment of peri-operative deaths in two hospitals. The authors commented that none of the range of hospital sources that held relevant data could provide all the eight required items. These were: patient name, date of birth, sex, hospital number, date of operation, date of death, consultant in charge and anaesthetist. Routine monitoring systems identified, at best, 58% of eligible surgical deaths and false positives were also common. In contrast, in the hospital in which coding clerks identified the deaths, complete and accurate data were obtained. In this hospital, coding occurred immediately after the death rather than after the discharge letters had been written. A trial of immediate coding in the first hospital improved results for 3 months and then ascertainment gradually dropped again. Cook and co-workers noted that an NCEPOD report had prompted changes in the seniority of clinical staff on duty, because this seemed to be associated with a greater risk of death but, in practice, found that it would have made no difference if senior staff had been on at night. They developed a risk model, but suggested that it was not specific enough to predict risk in the clinical setting. In 1997, 100 consultant anaesthetists were surveyed to ascertain whether or not their clinical practice was influenced by the NCEPOD; 74% said it was and 80% of anaesthetists found the NCEPOD feedback helpful. In 1998, Jelley, reporting on an NCEPOD briefing meeting, noted the difficulty of obtaining case notes for reviewing deaths, and the poor quality of the notes when found. This echoed the original conclusions of the CEPOD, which observed that it was up to consultants to ensure that their coding and input to monitoring systems was accurate and up to date.

Clark and co-workers calculated the sensitivity and specificity of the ascertainment of peri-operative deaths in two hospitals. The authors commented that none of the range of hospital sources that held relevant data could provide all the eight required items. These were: patient name, date of birth, sex, hospital number, date of operation, date of death, consultant in charge and anaesthetist. Routine monitoring systems identified, at best, 58% of eligible surgical deaths and false positives were also common. In contrast, in the hospital in which coding clerks identified the deaths, complete and accurate data were obtained. In this hospital, coding occurred immediately after the death rather than after the discharge letters had been written. A trial of immediate coding in the first hospital improved results for 3 months and then ascertainment gradually dropped again. Cook and co-workers noted that an NCEPOD report had prompted changes in the seniority of clinical staff on duty, because this seemed to be associated with a greater risk of death but, in practice, found that it would have made no difference if senior staff had been on at night. They developed a risk model, but suggested that it was not specific enough to predict risk in the clinical setting. In 1997, 100 consultant anaesthetists were surveyed to ascertain whether or not their clinical practice was influenced by the NCEPOD; 74% said it was and 80% of anaesthetists found the NCEPOD feedback helpful. In 1998, Jelley, reporting on an NCEPOD briefing meeting, noted the difficulty of obtaining case notes for reviewing deaths, and the poor quality of the notes when found. This echoed the original conclusions of the CEPOD, which observed that it was up to consultants to ensure that their coding and input to monitoring systems was accurate and up to date.

Lunn, the anaesthetist in the original study, CEPOD, published a critique of the NCEPOD in 1998 and noted a number of important developments. First, the NCEPOD no longer attributes a cause, but rather highlights cases for detailed study and aggregate analysis. It is still not possible to calculate case-fatality rates, because the Department of Health data collection does not record numbers of operations, and therefore the
review remains at the level of identifying remediable factors. However, while blame cannot be attributed, ‘rhetorical questions’ are asked to stimulate the readers of the report to examine their own practice. Second, each annual sample uses a uniquely devised questionnaire, which makes retrospective comparisons difficult. Moreover, the destruction of all computer tapes and questionnaires as part of the confidentiality agreement also reduces the potential for comparison. Thirdly, in respect of the NCEPOD’s achievements, Lunn noted that: audit meetings are now the norm; risk assessment by the NCEPOD has influenced the use of anaesthetic monitoring instruments; and records have improved. However, he recorded that the absence of a denominator is a crucial gap in the data, that there is no information on out-of-hours operations, and that there remains a danger that recommendations that are based on a sample of deaths may not apply to the care of the 99% of patients who survive.17

In 1998, Clinical Accountability, Service Planning and Evaluation carried out an evaluation of the NCEPOD.349 They noted its success as judged by the fact that 75% of eligible surgeons and anaesthetists participated in recording perioperative deaths, but made a number of recommendations for improvement:

• It would be helpful to put the level of morbidity in the context of the overall level of surgery undertaken in the participating hospitals and its relative success, so that the public could see the approximate levels of ‘background’ mortality.
• Local reporters should follow a single, nationally agreed, procedure for ascertaining deaths.
• The possibility should be explored of documenting the methods of case review and the practicalities of identifying specific criteria to be used in each review.
• An information leaflet would aid communication, both with local reporters and with clinicians in general, and might be valuable for other local staff involved in the care of patients.
• The resource implications of feedback to individual clinicians should be looked at because of the potential educational value.
• The NCEPOD should remain confidential, but consider retaining anonymous data to allow longitudinal studies.
• Especially with the advent of clinical governance, a strategy should be developed for following through on the NCEPOD’s recommendations to encourage their implementation.

In an important response, the NCEPOD Trustees re-emphasised the fact that the NCEPOD is not about the causes of individual deaths, but rather about general lessons to be learned from this group of patients.548,556 For this reason they feel that there is less need for denominator data, although it would be interesting. However, the absence of a denominator is a crucial deficit because the pattern of presumed risk factors and behaviours cannot be compared to that in a control or surviving population. In addition, core questions are now repeated for each annual sample, and these may allow some continuity. When reviewing CEPOD in 1987, Lunn and co-workers17 noted the need to develop rapid individual feedback without breaching surgeon confidentiality; this appears still not to be possible and, indeed, not to be an aim.

In summary, and against the criteria used for this review, the NCEPOD performs well as a system designed to identify factors associated epidemiologically with perioperative deaths, especially in respect of the special studies, in that it is judged by its participants to have contributed to stimulating critical audit and improvement of practice. However, as a monitoring system, which it is not intended to be, it has four main drawbacks: first, it does not include postdischarge deaths; second, there is no denominator to allow calculation of a case-fatality rate; third, the reports are not available until 18 months after the events; and, fourthly, it is a voluntary process with less than 80% coverage.

The Scottish Audit of Surgical Mortality

The SASM14 is one of a raft of ongoing audit projects funded by the Scottish Clinical Resource and Audit Group of the Scottish Office in 1993. Most projects were based on existing audit projects initiated by clinicians, and they have followed a broadly similar system of local audit coordinators and national collation of data and coordination of standard setting. The SASM resulted from the amalgamation of two audits of surgical mortality coordinated by the Royal Colleges of Surgeons of Edinburgh and Surgeons and Physicians of Glasgow. It was originally funded by a grant from the Clinical Resource and Audit Group for the three years 1994–1996, and since 1999 the administrative activity has been incorporated in the Information and Statistics Division (Edinburgh) of the National Health Service Scotland.

The Executive Group of the SASM comprises the presidents of the two surgical Royal Colleges in Scotland and a member of the Management...
Surgical mortality monitoring systems

Group from each of four regional centres. The Management Group consists of full-time administrative staff and the surgical and anaesthetic coordinators from each of the four centres. This group is responsible for the day-to-day running of the audit. The Advisory Group consists of members from all surgical specialties and anaesthesics on a geographic representational basis.

The SASM is a confidential audit of all deaths occurring in hospital under the care of a surgeon within 30 days of an operation. As with the NCEPOD, with the exception of obstetric and cardiac surgery, all surgical specialties in NHS hospitals are included, but it is not linked to the national hospital discharge data set. Although there is nearly 100% coverage of operations and procedures in the national Core Patient Profile Information in Scottish Hospitals (COPPISH) health information data set, it is not possible to match this accurately with the SASM data. There are, therefore, no baseline data from which to calculate operation-specific case-fatality rates, although a background rate can be calculated from COPPISH.

In 1998, 99% of surgeons and anaesthetists participated in this voluntary, professional audit. Data and relevant assessments were complete for 95% of all deaths by the deadline for data entry of 1 June 1999. Participation is considered to be a high priority by the Colleges in Scotland, and poor or non-compliance by individuals now generates a letter from the Chairman of the Executive Group encouraging more active participation. Staff in the local SASM offices are responsible for identifying all deaths in hospital or within 30 days of surgery, and do so with a network of contacts including medical records, mortuaries, secretaries and wards. Once a death has been identified and confirmed, data collection forms are sent, one to the responsible surgical consultant and one to the appropriate anaesthetist. The completed forms are returned in pre-addressed envelopes to the relevant SASM office.

At the SASM office, all identifiers are removed from the forms, which are then reviewed by an appropriate surgeon or anaesthetist. If there are inadequate data on which to conclude a preliminary assessment, the case records can be requested for clarification. In any case in which the preliminary assessment suggests factors meriting further scrutiny, the original forms and case records are sent to a second-line assessor in an appropriate specialty from another area of Scotland. The assessor completes a form and provides a short case summary and comment on the death. Approximately 10% of all deaths are subject to detailed case-note assessment. The data collected, including the commentaries, are entered in a customised database. The diagnoses and operations are coded using the NHS Clinical Terms (formerly Read Codes) and the adverse factors using a custom coding system developed by the SASM. It is impossible to make the case records anonymous, and therefore the audit is confidential. However, once the audit process is complete the data are made anonymous but not destroyed and can be accessed at the level of individual surgeon, surgical audit or Scotland.

Clinicians have the option of requesting feedback on all assessments but, by default, feedback is provided only when there has been a second-line assessment. Clinicians have the opportunity to respond to comments. The collected anonymous case commentaries are collated and circulated periodically within the profession. Individual hospitals or departments within hospitals can request reports and analyses of their data. A report is published annually on data collected in the preceding year and circulated to all participants. This comprises a statistical analysis in the forms of tables and figures with invited commentaries on each of the specialties. The latter are derived from a review of the collected data for the specialty, supplemented by more detailed analyses, as requested by the reviewer. The final draft of the report is approved by the Advisory Group before publication. It is estimated that the cost of the SASM is about £40 per death. There is some anecdotal evidence that the 1997 report produced a response in respect of health service provision.

Eight publications were reviewed that dealt directly with the SASM or the Lothian Surgical Audit, which was one of the predecessors of the SASM. The primary assumption of the Lothian Surgical Audit was that the national Scottish Hospital in-patient statistics were not suitable for the purpose of surgical audit because they were inaccurate, contained insufficient detail of operations, and were too late in appearing, which meant that the staff concerned or the organisation of services had changed in the interim. In 1989, Clason and co-workers showed that the data, augmented by prospective recording of all operative procedures, could be used by a vascular specialty to make comparisons.
of morbidity and mortality over time. Bradbury and co-workers\textsuperscript{360} also augmented the Lothian Surgical Audit in one hospital to create ‘morbidity profiles’ (i.e. complications). In 1997, Aitken and co-workers\textsuperscript{361} reviewed the Lothian Surgical Audit and concluded that it had changed practice in respect of the establishment of specialist units, improved results after bowel surgery and led to the demise of outmoded techniques. They believed that this success derived from the facts that the audit was driven by surgeons, could be flexibly used in everyday practice and had inspired confidence in participating surgeons. In 1998, MacArthur and co-workers\textsuperscript{362} reported on the Scottish Mortality Study, also a forerunner of the SASM. In a nested confidential, but not anonymous, audit the case notes and audit booklet were sent to an independent assessor and then returned, with the assessment, to the consultant concerned. Anastomotic leak was the single most important adverse event, and many of these events were judged avoidable. Finally, in 1999, Stonebridge and co-workers\textsuperscript{16} examined participation by the 1000 surgeons within the SASM and reported a steady 93\% over its first 3 years.

In conclusion, the SASM is a comprehensive national audit of surgical mortality. There are problems with its structure, which include its dependence on the quality of the case record on which the assessor judges the presence of adverse factors in outcome. Furthermore, the lack of a denominator in terms of the numbers of patients treated means that the magnitude of any observation identified by the SASM cannot be placed in context. However, its recent incorporation into the Information and Statistics Division may facilitate this process in the future.

Monitoring systems within other surgical specialties

Four further examples of nationally coordinated continuing audits serve to demonstrate the application of the approach to a range of topics.

The Scottish Trauma Audit Group

The Scottish Trauma Audit Group (STAG) has a central office and local coordinators in each of the 15 health boards.\textsuperscript{363} It was originally funded by the Clinical Resource and Audit Group in 1991, but is now funded jointly by Scottish Health Boards. The local coordinators are nurses who are trained and supervised by one of three regional coordinators for Scotland. Data are collected on inpatients who are admitted for at least 3 days, or who die in hospital, except for patients aged under 13 years and those over 65 years who have sustained an isolated fracture of neck of femur. The event of primary concern is inpatient mortality with no specified time frame. There is no population denominator or postdischarge coverage, but inpatient case-fatality rates can be calculated for the registered patients. Coverage is 25 hospitals dealing with approximately 98\% of seriously injured patients in Scotland. There were 6290 notifications in 1998.

There is a strong emphasis on ensuring completeness and accuracy. Data collection is started by the accident and emergency nursing staff, who provide physiological and time-based data. The local coordinator is responsible for extracting data about prehospital and in-hospital care from the case notes. The forms are made anonymous and sent to the national quality assurance manager for checking, where additional information is sought if necessary. There is central computerised validation for inconsistencies, and each hospital carries out a daily check for patients who have been missed. The STAG uses the revised Trauma Score and Injury Severity Score (TRISS) to quantify severity for the calculation of expected mortality. In 1997–1998 the proportion of eligible patients entered was 94\% (84–100\% interhospital range). The local coordinators are trained and supervised, and there are more than 200 computerised validations. In 1997 the STAG was awarded ISO 9001, the first audit project to achieve this quality assurance standard, and in Scotland it is regarded as the gold standard for audit projects.

Feedback and output are at three levels:

- monthly data summaries are sent to each participating hospital
• a 6-monthly analysis comparing an individual hospital with the rest of Scotland is sent to each local medical director and coordinator
• an annual report is sent to each Director of Public Health.

In addition, results of associated research projects are distributed to participating accident and emergency departments, and there is an ad hoc data management service for those who wish to conduct research. To date, the STAG has published 14 papers in peer-reviewed journals. The visible influences of the STAG on care include the development of a neurosurgical referral letter to improve the process of referral.

The UK Trauma Audit and Research Network
A parallel system to the STAG is the UK Trauma Audit and Research Network (UKTARN), which runs in England, Wales and Northern Ireland. It covers almost half the relevant hospitals of England and Wales, and a smaller proportion in Ireland. It follows similar data collection procedures to the STAG, with specially trained staff. Outcome, in terms of survival or death, is assessed at discharge or 3 months, whichever is first. Re-analysis of records showed 90% agreement of abbreviated injury scale coding, and an inter-rater reliability of 97% when the scores were used to calculate the injury severity score. Thus the main reported difference between the STAG and this trauma system is the level of overall coverage, and the high inter-hospital variation in reporting in the UKTARN.

The Scottish Hip Fracture Audit
The Scottish Hip Fracture Audit (SHFA) was initially funded by the Clinical Resource and Audit Group in 1993, but the cost is now shared by local participating health boards. The SHFA covers approximately 80% of the 6000 hip fractures occurring in Scotland each year (SHFA, personal communication). The system is very similar to that of the STAG, but mortality is only one of the outcomes of interest and the focus is on improving care in terms of both quality and efficiency. Mortality is defined as death up to 120 days postfracture, regardless of location. Standard data collection protocols are applied at admission, at 4 and 12 months and at any hip-related re-admission, with reported coverage of follow-up of 98%.

Data collection is conducted by dedicated audit nurses and the standard core data set includes case mix, surgical procedure, complications (e.g. pressure sores, wound infections), mobility, dependency, residential status and mortality. Items of local interest may be added. Unspecified validity checks are made for accuracy and completeness in each hospital. Feedback to the local centres comprises 6-monthly detailed and individualised reports. An annual comparative report and published report is planned for this year. The hip fracture audit is popular because it is distributed and combines national comparisons with local topics of interest. It also generates much research activity and links with Europe to develop standardisation of ongoing audit across Europe. To date, there is no published information on its impact on care.

The Scottish Hip Arthroplasty Register is also mentioned here because it is an example of combined generic and special interest data collection. The national general hospital discharge form (SMR01), which records the main condition treated, complications and co-morbidities, and up to four procedures, is augmented by extra clinical data and linked to Registrar General death data. This system has been used to ascertain the incidence of DVT after arthroplasty in comparison with that after cataract surgery.

The Royal College of Surgeons of England Comparative Audit Service
In 1990 the Royal College of Surgeons of England set up a voluntary confidential Comparative Audit Service so that its general surgical fellows could compare their results confidentially. This service continues to operate within the Clinical Effectiveness Unit, formerly the Surgical Epidemiology and Audit Unit. The data collected include surgical workload, case mix, number of operations performed, deaths and complications. The aim is to encourage local systematic and standardised continuing audit by providing a national perspective for valid standard setting and comparison. The college has supported large-scale audits in ten topics: colorectal cancer; hernia surgery (National Groin Hernia Outcomes Project); prostate surgery (National Prostatectomy Audit); total hip replacement (National Audit of Total Hip Replacements in conjunction with the British Orthopaedic Association); cleft lip and palate; ankle fractures; gastrointestinal bleeding; intrathoracic transplantation (UK Cardiothoracic Transplant Audit); liver transplant (UK Liver Transplant Audit); and patient satisfaction. Several of these large-scale audits have been single projects, but some are ongoing. Mortality
is defined as deaths from all causes, inpatient only, and case mix or risk is calculated using age, admission status and diagnosis. The main outcome measure is morbidity, and complications are classified by three grades of severity (major, intermediate, minor). The output is comparisons that have been made anonymous, but there are no annual reports. However, many of the audits have been published in peer-reviewed journals and have been used for the development of college guidelines, such as for the management of colorectal cancer and gastrointestinal haemorrhage.

Cardiac surgery monitoring systems

This section describes the procedure- or specialty-specific surgical monitoring systems within the UK and the USA. The main examples are cardiothoracic and vascular surgery. In the UK, three databases are maintained by the Society of Cardiothoracic Surgeons of Great Britain and Ireland (SCTS). In the USA, the main adult cardiac surgery database is run by the Society of Thoracic Surgeons (STS), although smaller cardiac databases and registers exist within certain States. Two national Scandinavian vascular surgery systems are also described.

Within the UK, the three programmes run by the SCTS are: the UK Cardiac Surgical Register, the National Adult Cardiac Surgical Database and the UK Heart Valve Registry. In 1998, the first two systems jointly covered 70% of cardiothoracic centres in the UK. One additional database, the Central Cardiac Audit Database, a combined medicine and surgery audit, is also currently being piloted.

The UK Cardiac Surgical Register

The UK Cardiac Surgical Register was established in 1977 by Sir Terence English. It is based on voluntary and anonymous reporting of activity and hospital mortality for all cardiac surgical procedures performed in NHS hospitals in the UK. It covers approximately 35,000 cardiac procedures per year. This venture was the “first attempt by any surgical or medical specialty to capture nationwide data”.

Mortality is defined as death within 30 days of surgery. Mortality data collection has been reported by Keogh as “difficult to track and validate”, and it has not been rigorously validated. There is no denominator for the calculation of population rates. Information on cardiac surgical procedures is collected by requesting each unit to complete a standard form that is returned to the SCTS, made anonymous and passed to an independent consultant for aggregation and analysis. The data collection form has two sections, one for adult and one for paediatric cardiac surgery. CABG surgery in adults is collated, as are coarctation and ventricular septal defect repair in children. For CABG, the surgeon-specific data include simply the number of operations, the mean Parsonnet (risk) score (if known) and the number of deaths. For children, definitions are given of which procedures to include. For thoracic surgeons, only numbers of lobectomies for cancer and numbers of deaths are submitted.

These unit-based data are then aggregated into an annual report that has provided useful information on cardiac surgical activity. Up until 1997–1998 only surgical units were identified, but since then the register has been extended to include activity and outcome data on individual surgeons for some specific operations, the purpose being to “help restore public confidence” in the specialty. The reports of the collated data in the register are sent to each member of the SCTS for their own use.

The UK Adult Cardiac Surgery Database

The Adult Cardiac Surgery Database is the most comprehensive of the three run by the SCTS. It was established in 1994, and represents a move from simple aggregated data to patient-specific data. In 1998, just over half of all British cardiac surgical units submitted data. There are precise and repeatable definitions for all the 150 or so items in the Minimum Dataset. It is important to note that, because of the recognised difficulties of ascertaining deaths after discharge from hospital, the definitions of death were changed from that in the Cardiac Surgical Register. It is now “death on the same admission as surgery in the base hospital”, and the SCTS notes that it not only excludes postdischarge deaths but it also makes the assumption that dying patients are not transferred to other hospitals in significant numbers. The definition also includes deaths during the same admission, but after 30 days, making it potentially different from most other registries.

The database includes risk factors, postoperative complications and interventions. The aim of the database is to develop reliable comparative UK stratification models in conjunction with the Medical Research Council Biostatistics Unit at Cambridge. The SCTS notes that the data are compatible with, but a reduced version of, the US STS data set. It is also compatible, in terms
of content and definitions, with the UK Heart Valve Registry and the European Cardiac Surgery Registry.

Data are collected from local systems to a central merged database that uses a patient analysis and tracking system. Source registries are entered in total and then converted to a common data structure to allow different analyses, including comparisons over time. There is data validation for logical errors before the patient analysis and tracking system registry is made anonymous (it holds unique identifiers only, not names). Each centre is sent their own now anonymous data for comparison with national analyses in order to permit the use of centrally produced risk models. As yet, however, there has been no external validation of completeness and accuracy.

The risk modelling yields Parsonnet and EuroSCORE scores, receiver operator characteristic curves and calibration plots to compare the sensitivity and specificity of different risk modelling techniques, all of which are still being refined. In 1998, Keogh reported that the two most robust outcomes are surgical mortality and hospital length of stay. A data validation in 1996 for four risk factors showed data completeness averaging from 60% to 88% for four factors. Only centres submitting an average of more than 90% were included in the risk analysis. The 1998 report was published in May 1999 and the website is regularly updated; thus feedback is timely and participants appear to like it. The costs of running the database have not been identified.

The UK Heart Valve Registry
The UK Heart Valve Registry was set up in 1986 and is funded by the Department of Health through the Medical Devices Agency. It monitors the valve rather than the operation, and records mortality at any point after implant by virtue of its linkage to the national death registers for the four countries of the UK. Thus it captures a true 30-day mortality and allows longitudinal analyses of death. Although the registry has historically collected data from NHS cardiac centres only, because of the overriding need to be able to identify all patients at risk after heart valve replacement surgery the registry now collects data from private cardiac surgical units throughout the UK. The death data comprise date, certified cause of death, including where, and any post mortem information that is available. A link person, not necessarily clinical, is identified in each collaborating centre as responsible for the data collection, and is given a set of guidelines for form completion. Reports for a participating centre can be supplied at any time, but not reports for other centres, although a collated anonymous annual report is published annually.

Although the coverage of the system is not reported, the linkage to death records should make this potentially a very tight monitoring system of adverse events associated with different heart valves. Actuarial survival can be calculated, and the reported intention is to publish 10-year follow-up interval data. A 98% complete follow-up over a 10-year period was reported in 1998. Furthermore, because the registry contains the number of each valve prosthesis, it fulfils a regulatory role.

US cardiac monitoring systems
Within the USA, the main adult cardiac surgery database is run by the STS. In addition, there are national and regional registries throughout the USA, including the New York Cardiac Surgery Reporting System, the Northern New England Cardiovascular Study Group and the Veterans Affairs Continuous Improvement in Cardiac Surgery.

It is important to note that cardiothoracic monitoring systems in the USA were developed largely in response to the furore when New York State published death rates in which first hospitals (1986) and then surgeons (1991) could be identified. At least initially, therefore, their prime purpose was risk prediction to defend case-fatality rates above the national norm. Increasingly, however, they are presented in the context of quality improvement and audit, an ethos which chimes more closely with the aims of UK and other European database registries. There are many intentional similarities between the various main registries (e.g. the SCTS in the UK is linked into the STS), but there are also some important differences. There is more of a focus on risk-adjustment and morbidity, not just on surgical mortality. The examples of monitoring systems given here are the ones that are widely quoted in the literature by the cardiothoracic experts, and describe their data collection and processing systems.

The STS database
The STS runs a Data Warehouse that contains the STS Adult Cardiac Surgery Database and a Congenital Surgery Database, covering both the USA and Europe, which is currently under
Voluntary data collection on cardiac procedures from a small number of institutions began in 1980, and these earlier data were used to create a risk model. Three separate databases are now held: the Adult Cardiac, Valvular and Congenital Heart Surgery databases. The STS databases are held at the Duke Clinical Research Institute, where analyses are conducted. By 1996 all but one of the US states was represented, although some large centres, such as Johns Hopkins, were still not involved. By 1999, at least one million patients were enrolled from about 70% of US cardiothoracic surgeons, covering all types of cardiothoracic surgery. The system remains voluntary, and the denominator for detection is all registered procedures.

In 1996, Clark, one of the originators, reviewed benchmarking issues. He noted that completeness, accuracy and external validation had clearly not been achieved, but that there was 95–98% completeness of items needed for risk stratification, with clinical features being the most poorly reported. Intra-record consistency checks were now in use, and it was thought that purposeful gaming (to overestimate risk, and therefore ‘explain’ high mortality rates) was less than 1% of records. He observed that the advent of data managers and the introduction of local audits meant that it was much less likely that gaming by surgeons would be possible or pass unnoticed. However, Shroyer and co-workers reported that 12% of CABG records had missing operative mortality data. Therefore, since then the Audit and Validation Sub-Committee has instituted new quality control measures:

- feedback to the hospital on the initial harvest of cases
- confidential letters and, if necessary, site visits for persistent problems
- enhanced software to permit intra- and inter-field checks of completeness and quality according to published criteria
- a national nurse coordinator to provide educational activities
- a new clinical issues enquiry service on the website.

In future, if more than 5% of a surgical group’s records have missing operative mortality information, the entire group’s data will be rejected.

Risk-adjusted 30-day mortality for CABG, non-risk-adjusted 30-day mortality for other cardiothoracic procedures, and hospital lengths of stay are reported annually and permit comparison of local with national results. Publication of the annual update of the risk model takes at least 2 years; the full results for 1998 are on the STS website. However, individual hospitals have their data returned earlier. There have been many audit and research publications using the database and, in a new development, the STS is working with the HCFA to compare outcomes for coronary bypass and angioplasty. In 1996, for example, Nikas and co-workers used the database as a baseline against which to monitor outcomes when they introduced a change of patient management.

In summary, although the focus of the STS database is on risk-adjusted mortality, it contains data relevant to other surgical adverse events. However, it is recognised that even the ascertainment of death is not complete and, in particular, that there is no system of identifying deaths after discharge from hospital.
The Cardiac Surgery Reporting System, New York

The Cardiac Surgery Reporting System is a database and risk-adjusted outcomes programme, started in 1989, that tracks all cardiac surgery, both adult and congenital, at all 31 cardiac centres in New York State.\(^{375}\) It has been developed and managed by the State Department of Health, with input from the Clinical Advisory Committee, a panel of consultant physicians.

Standard definitions are used and mortality is defined as ‘in-hospital’, regardless of timing. Approximately 40 items of data are collected, including risk factors, postoperative complications and hospital and surgeon identifiers. Data forms are completed in cardiac surgery departments and transferred to a database system on a personal computer in each hospital. The data are then forwarded quarterly to the State Department of Health for processing and analysis. A utilisation review agent examines a sample of medical records for accuracy of coded data, and hospitals are asked to re-submit data for any problems that are identified. There are occasional site visits. Surgeon-specific data are released only for surgeons who do more than 200 operations in a 3-year period.\(^{382}\) In 1994, Hannan and co-workers\(^{375}\) reported that the Cardiac Surgery Reporting System was better than the New York Statewide Planning and Research Co-operative System in predicting case-specific mortality, because of the inclusion of clinical detail. In a further study, Hannan and co-workers\(^{385}\) linked prospectively defined data for CABG and angioplasty to the state’s vital statistics to compare 3-year outcomes of revascularisation and death.

The Northern New England Consortium

The Northern New England Consortium was established in 1987 to represent all institutions performing cardiac surgery.\(^{382}\) Today it is the Northern New England Cardiovascular Disease Study Group and comprises open heart surgeons from six institutions.\(^{384}\) It is a voluntary system and there is self-selection of participants. Definition of mortality is defined as status on discharge (i.e. in-hospital mortality only). These registries cover more than 60,000 patients, with an annual accrual of approximately 8000 patients. Practice on data collection varies by hospital and there is a confidential web page with clear guidance and definitions for submission of data.\(^{375}\) From the beginning the focus has been on improving quality of care and, within the consortium, the Dartmouth Group has developed data collection on cost-effectiveness and patient satisfaction, using patient focus groups and functional health status measurements.\(^{385}\) It has also developed a programme to calculate individual risk at the patient’s bedside,\(^{385}\) although no validity or reliability studies of this programme have been reported. This study group has a clear output and utilisation strategy of continued and regular feedback of outcome data, but no publication of impact was found. This is an organised effort to understand better the processes contributing to mortality. In 1999, the central costs of the system were US $300,000 per annum, or approximately US $30 per entered chart, which was less than 1% of the total cost of a CABG.\(^{384}\) Local costs were also estimated at about 50 cents per chart.

The Department of Veterans Affairs

The Veterans Affairs Cardiac Surgery Consultants Committee was formed in 1972 to monitor hospital performance.\(^{386}\) In the mid-80s they moved to risk-adjusted data using a single-page data collection form that included preoperative and operative variables for all patients undergoing cardiac or great vessel procedures requiring cardiopulmonary bypass, and also outcome data including ten complications. The Veterans Affairs Continuous Improvement in Cardiac Surgery programme is run from the centre at Denver, Colorado.

Mortality is defined as any death within 30 days postoperatively and any postoperative death directly related to the operation regardless of location. Data collection is mandatory from all Veterans Affairs hospitals, and therefore provides a denominator of all such operations. If more than 20% of data are missing the variable is not included, and cases with missing values are not used.\(^{387}\) Inter-rater reliability is checked periodically in the system by travelling nurses.\(^{387}\) Data are analysed separately for CABG patients and for those undergoing valve or other non-CABG procedures.\(^{388}\) There is 6-monthly feedback of unadjusted and estimated risk-adjusted operative mortality to individual hospitals, although the only identifiable hospital is the recipient’s. In addition to the 6-monthly ‘real-time’ feedback, adjusted 3-year trend data are produced. In 1996, Grover and co-workers\(^{386}\) commented that the tracking of postoperative morbidity was less successful than of mortality and that the Veterans Affairs system still lacked uniform definitions. Furthermore, much of the data were missing and the reliability of the various risk models within the Veterans Affairs system required validation.\(^{386}\) In April 1997, new, explicit definitions and criteria for data collection
were introduced, covering demographics, clinical catheterisation, angiography, operative risk, operative outcome and resource-use data sets. In 1994, Grover and co-workers reported a wide range of clinical and structural deficiencies that had been highlighted and, mainly, addressed locally after receipt of the feedback and site visits. Clinical problems included poor communications, operating technique and inadequate supervision, while structural ones were inadequate equipment and insufficient nurse/patient ratios in the intensive care unit.

The National Veterans Affairs Surgical Quality Improvement Program was established in 1994. The focus moved to one of continuous monitoring and enhancement for all surgery, rather than simply risk adjustment for cardiothoracic surgery. This programme is described below under the heading ‘generic monitoring systems’.

In 1994, Daley critiqued five programmes that monitored risk-adjusted outcomes in CABG. These were the New York State, Northern New England, Parsonnet, STS and Veterans Affairs programmes. She first noted that, in terms of definition, only the Veterans Affairs Continuous Improvement in Cardiac Surgery included deaths up to 30 days regardless of where the death occurred. All the other monitoring systems only included deaths in the base hospital. Second, only the Veterans Affairs monitored morbidity as well as mortality. Third, she concluded that the precision of the standardised and published definitions for identification, classification and recording of risk factors was high in most US and UK systems. Lastly, the New York State, STS and Veterans Affairs systems had assessed data accuracy and completeness by comparing samples of charts against submitted entries. Validation was conducted by external reviewers (STS, New York) or paid data managers (Veterans Affairs), both to assess reliability and to detect any ‘gaming’.

In summary, the prime purpose of the US monitoring systems has been risk adjustment to explain higher than average case-fatality rates (i.e. the focus has been on attribution rather than on monitoring for improvement). This is now changing. Morbidity is seldom looked at except as reflected in re-operation and it is not clear that the systems can detect, or are interested in, very rare events. Outputs are in the form of risk models based on rates of occurrence and statistical handling. From this literature, descriptions of impact are scarce.

Canadian cardiac monitoring systems

One Canadian example of a cardiac surgery monitoring system was found. The British Columbia Provincial Cardiac Registry was created in 1990 in response to unacceptable waiting lists for cardiac surgery, such that patients had to be referred to the USA. This surgical monitoring system covers the whole province and is based on a system that originated in Vancouver General Hospital. It comprises cardiac surgery, angioplasty and pacemaker/defibrillator subsets. The Cardiac Surgery Database includes cardiac surgery booking, operative report, waiting list reconciliation and a discharge summary, and is used for future clinical care, as well as for education and research. Surgeons complete a 16-page form at the time of operation, which a clerk then enters into the database, and a printed report is stored in the patient’s record. Selected data fields are uploaded nightly to a central provincial database.

The denominator is all procedures. A detailed validation study for reliability compared the database with hospital charts and found an inconsistency rate of 9.9%, which was traced to surgeons rather than the clerks. There was no relationship between the volume of operations per surgeon and consistency of reporting. The most accurate items were those collected at the time of operation, and clinical information was the least reliable. Fields that were captured by check-off boxes, rather than multiple-choice or fill-in-the-blank methods, were the least reliable. The authors noted that a request for an actual number was more consistent than a subjective interpretation, and that accuracy appeared to be encouraged by providing definitions and recording specific values rather than ranges. The authors recommended changes to the system based on their findings, starting with feedback to users on the aims of the system and its performance. There were no published reports of its impact.

Vascular surgery monitoring systems

The Swedish Vascular Registry

The Swedish Vascular Registry (SWEDVASC) was started in 1987 as the Vascular Registry in Southern Sweden. By 1990, half the country was covered and the name was changed to SWEDVASC. In 1998, Bergqvist and co-workers reported that the database now covers vascular surgery in all types of hospitals in Sweden, so that over 90% of the population and of elective major procedures are included. Cases overlooked are mostly minor procedures and some re-operations. Mortality
Surgical mortality monitoring systems

is defined as up to 1 year, but deaths continue to be notified thereafter by linkage from the national population registry. A denominator exists for both population- and case-based rates. Approximately 30 risk factors, operative events, postoperative events and assessments are recorded. The decision on whether a complication, such as wound infection or stroke, is eligible for inclusion is left to the surgeon’s discretion.

Data collection is conducted by surgeons who complete a form at 30 days and 1 year postoperatively. These reviews are prompted by monthly lists from the register. The completed forms are mailed to a secretary, and reach SWEDVASC via the Cancer Registry. Identified events are taken back to hospital level and patients’ records analysed and coded. This process complies with patient privacy requirements.

Several publications have reported on the process of data validation. In 1989, Troeng described re-registration of a 10% sample by the responsible surgeon after the first 6 months of the registry. A second validation study compared registrations against operation lists and anaesthesia registries. Re-registration showed a 90–95% agreement for most variables, but less consistency for preoperative risk factors. Coverage was shown to be 90–100%.

In a study of intestinal ischaemia after vascular aortofemoral surgery, Bjorck and co-workers reported that reproducibility checks of random samples of cases were performed regularly and confirmed a 90% report rate and 90% reproducibility rate. A further 5% sample of non-ischaemia operations was identified and records checked. Only three further cases were found, all of whom were ill for other reasons and died. The authors suggested that for total ascertainment a special check must be made among early non-survivors, in whom autopsy is likely. In 1994, Bergqvist and co-workers noted a decline in 1-year follow-up and the 1-year value for the ankle–brachial index was missing in 30% of the cases reported. The registry was used to identify re-operations and revisional surgery over a period of 4 to 9 years. This study found no simple or consistent explanation for the repeat interventions and few risk factors. Ascertainment is more likely when clinical symptoms are present and in county rather than university hospitals.

Bergqvist and co-workers, reviewing the first 10 years of SWEDVASC as an outcome study, observed that vascular surgeons have not agreed on common standards of quality control, but that the concept of sentinel events was important in agreeing which variations are acceptable. In SWEDVASC, four such events were sought:

- re-operation for occlusion in femoropopliteal bypass surgery (possible technical error)
- amputation within 30 days after operation for claudication
- stroke within 30 days after carotid endarterectomy for transient ischaemic attacks
- graft infection after elective procedures.

They found a low and constant frequency of 1–2% of these complications. However, they noted that it was not possible to assess the influence of the registry on improving care.

In respect of output, there is a yearly publication with aggregated data for common use, and hospital-specific data for local use. Reported adverse events are re-operation, graft infection and death. The registry has been used for a range of research projects that would not otherwise be possible. It is estimated that the cost of the registry (1998) is approximately US $6 per case.

The Finnish Vascular Surgery Registry

The Finnish Vascular Surgery Registry (FINNVASC) was established in 1991, using a record form modified from SWEDVASC and similar data collection procedures and definitions. However, FINNVASC is not linked to national death registration. It covers the whole population of Finland but omits three private and 18 small district hospitals. Validation of internal accuracy and completeness was carried out both initially and in 1997. Cross-validation against computerised hospital records found a mean of 19% missing cases (0–47%), and these were most likely to be emergencies and endovascular procedures. When 2% of forms were refilled, there was a 93% agreement (81–100%), but less than half contained no differences.

In 1998, the Steering Committee of SWEDVASC reported an increasing interest in the development of vascular registries and mentioned The Netherlands, Northern Ireland, New Zealand and the European Community. However, no further published literature was identified.
Generic UK national monitoring systems

The previous section described the process of monitoring surgical mortality within dedicated systems. In the UK there has been considerable interest in the use of national data and the four UK systems were reviewed for their potential to monitor surgical mortality.

Hospital Episode Statistics, England

The English Hospital Episode Statistics system has been running since 1988–1990 and is based on records relating to each patient treated as an inpatient or day-case in NHS hospitals. Before this time, the Hospital Activity Analysis covered similar data from 1968. Inpatient episodes are linked within and across hospitals to comprise a continuous inpatient spell. Patient records ‘flow’ electronically within a complex network in a common data set standard from the patient administration system from each trust.

In the context of surgical mortality, the assessment of deaths of hospital inpatients is up to 30 days, and the denominator is “all surgical continuous inpatient spells”. The main output of the English system is the feedback to trusts of data quality indicators and a conference for open discussion of data flows and products. The age-standardised data are used in a wide range of national reports and are also produced as national clinical indicators for a range of operative procedures. The indicators have not yet been assessed for their sensitivity, specificity, repeatability and responsiveness to change.

COPPISH, Scotland

This scheme, previously called the Scottish Morbidity Record (SMR) scheme, started in 1957 and covers 100% of discharges in all types of Scottish NHS hospitals. Surgical information is derived from SMR01 for inpatients and day-cases; outpatient coverage (SMR00) has started but is not yet complete. Additional information about cancer patients is available from cancer registration, which includes automatic linkage with national death records from the General Register Office for Scotland.

Deaths can be presented for any time postoperatively, but the customary time for ascertainment is 30-day mortality, obtained by special linkage with the General Register Office for Scotland. Diagnoses and procedures should be coded by medical staff using ICD and OPCS codes. Up to five complications can be identified by the use of the relevant ICD codes, but the form does not distinguish between co-morbidity and postoperative complications. Data are then extracted by clerks from hospital discharge letters and case records. The Information and Services Division (Edinburgh) runs training and operates a help desk for coders, as do most of the national systems.

With regard to validation, there are 100% logical error checks included in the Information and Statistics Division software. Additional validation studies have been conducted on the COPPISH–SMR system. In 1994 and 1996, 10% samples of the validated discharges for a 1-month period found, by comparison with case records, that the accuracy of ‘main diagnosis’ was greater than 89% and that of ‘other diagnoses’ over 90%. Completeness of ‘other diagnoses’ was 85%.

Recording of operations and procedures achieved an accuracy of 94% and 93%, respectively, and completeness of 97% and 90%, respectively.

When analysed by hospital, the levels of both accuracy and completeness ranged from 88% to 93%. However, in 1996, Harley and Jones noted a high level of errors due to omitted secondary diagnoses and procedures.

A recent pilot study looked at the potential to use national routine data for clinical audit and benchmarking rather than parallel special data collections. It noted that, for this purpose, the current data set needs to be supplemented by patient-group-specific information. Crucially, it also reiterated a frequently expressed view: namely, that unless clinicians themselves collect and record the required items in an accessible part of the record, ascertainment will continue to be less than perfect. Evidence on the user-friendliness of COPPISH and the use of the data to improve care is scarce. A recent in-house study reported that clinicians do not trust the accuracy and completeness of the SMR01 data and, if interested in audit, may still be inclined to create their own additional system. It is hoped that adherence to the use of a recent Scottish Intercollegiate Network guideline on a nationally
agreed hospital discharge form will be the trigger for improvement in data completeness in the national system.

The denominator for monitoring surgical deaths is all discharges from NHS hospitals. In Scotland, less than 10% of surgery is conducted in private hospitals. Initially, feedback of annual activities compared to the national norm was given to all hospital consultants. This has been replaced by open access to the Scottish Key Indicators Package for Performance (SKIPPER) database, which is now available on the Internet. Summaries of the data are also published within 1 year as Scottish Health Statistics. The Scottish Clinical Outcome Indicators, which have been published annually since 1992, have been used as the basis for local enquiry into the levels of case fatality for a range of surgical procedures and conditions. There are no recent published data on costs.

The Patient Episode Database for Wales
The Patient Episode Database for Wales (PEDW) was introduced in April 1991 to replace the Hospital Activity Analysis and the Mental Health Enquiry. It contains data on all inpatient and day-case activity undertaken in the NHS in Wales. Data are also received from Cleenet on Welsh residents treated in English hospitals. Finance for the data set is supplied by the National Assembly for Wales Health Information Management & Technology Division (HIMT), via Health Solutions Wales.

There is no link to deaths in the community, and therefore PEDW contains information only on deaths that occur in hospital, regardless of timing. Data held on complications depend on whether the relevant ICD-9 or ICD-10 diagnosis code is recorded but, as with the other UK systems, comorbidity cannot be distinguished from complications. Similar data-cleansing rules as those applied to the Hospital Episode Statistics are used on data held on the Warehouse.

The PEDW data are used for the same range of national reports as in England and Scotland. Data quality indicators for the PEDW are published on Cymruweb (intranet) and are updated on a monthly basis. Examples of the indicators produced are the percentage of episodes with: a valid principal diagnosis, a valid postcode and a valid registered GP. There is, however, limited validation at the source of the PEDW extract. An information quality programme for the NHS in Wales is currently being introduced and includes the development of a data quality indicator.

Northern Ireland
Little information was retrieved on routine statistics in Northern Ireland. The system for validating and using hospital inpatient data in Northern Ireland is less well developed, although this is likely to change in the near future.

Generic US national and regional monitoring systems
All major US health organisations and funders have routine data collection systems for billing purposes. This requires a level of detail and accuracy not available in the UK national systems, and the main US sources were examined to see if they would yield systematic information about surgical adverse events.

Health Care Finance Administration
The HCFA holds the administrative records of the Medicare and Medicaid programmes. In 1984 it created Peer Review Organisations to use claims data to monitor the costs and quality of care under Medicare. The Healthcare Quality Improvement Programme began in 1996 and covers a wide range of topics. The new HCFA approach emphasises positive improvement, rather than the earlier punitive ethos and the defensive approach that it produced, by which the main aim was to demonstrate that variations in death rates were usually attributable to differences in the risk scores of the patients. The prototype of the new approach, which is more like that of the UK and Scandinavian surgical monitoring systems in looking at patterns rather than cases, is the HCFA Cooperative Cardiovascular Project. The data sets of the HCFA have been the subject of most of the published papers on generic versus dedicated data sets for monitoring purposes; these are discussed in chapter 9.

The Department of Veterans Affairs
As well as the Veterans Affairs Continuous Improvement in Cardiac Surgery described above for cardiac surgery, the Veterans Affairs holds a mandatory database of all patients and treatments in Veterans Affairs hospitals. From 1991 to 1993, the National Veterans Affairs Surgical Risk Study was run in 44 Veterans Affairs Medical Centers and was used to develop risk-adjustment models for 30-day mortality and morbidity rates for all non-cardiac surgery. Separate models were created for general surgery, vascular surgery, orthopaedic surgery, urology, plastic surgery, thoracic (non-cardiac) surgery, neurosurgery and otolaryngology. These models were
validated and then used to establish the National Veterans Affairs Surgical Quality Improvement Program in January 1994. This now acts as a reporting and managerial structure for the continuous monitoring and quality enhancement of surgical care in Veterans Affairs medical centres. Surgical mortality is defined as the death of a patient within 30 days after the index surgical procedure, in or out of the hospital. Surgical morbidity is defined as the presence of one or more of 20 predefined complications in the same time period. The denominator is all operations in the participating specialties and hospitals. About 100,000 major surgical procedures are reviewed each year.

There are ten teams of trained surgical clinical nurses across the 123 medical centres involved in the programme, which ensure the accurate collection and timely transmission of data. There is a standard operations manual and definitions, and data are collected on workload and for risk adjustment. The latter comprise 45 presurgical factors, 17 surgical factors and 33 factors associated with outcomes, including complications.387 A computer-generated follow-up letter requesting information is sent to every patient 30 days postoperatively. At 45 days, the nurse completes the patient’s data entry in the risk-assessment module; it is agreed by the chief of surgery and transmitted to the coordinating centre (the same procedure is used for both non-cardiac and cardiac surgery). Laboratory data are transmitted automatically. At the coordinating centre the data are edited for missing, out of range and inconsistent entries. Records with potential errors are checked at the originating centre. An audit trail of all data corrections is kept. All deaths are verified against the Veterans Health Administration Beneficiary Identification and Records Locator System death records. A monthly inventory is sent to each centre for the nurse to check for completeness.

Feedback on hospital rates of death and complications is primarily through an annual evaluation by peers of the results at each medical centre, and in an annual report that goes to all the local chiefs of surgery and the hospital. This permits comparison of local versus national performance, but only for one’s own hospital. The Veterans Affairs Continuous Improvement in Cardiac Surgery has an additional semi-annual report. The database can be used for research that is approved by the Executive Committee.

Data on the total surgical volume in the Veterans Affairs was available for 1997. Of 343,808 surgical operations, 157,226 were classified as major. Of the latter, more than 30% were ineligible for assessment in the system because of a lack of nurses to review them or because of excluded procedure codes. However, 97.7% of the eligible cases were assessed. The cost in the same year was assessed as around US $12 per surgical procedure and US $38 per major surgical procedure assessed by the programme.

In 1998, Khuri and co-workers387 warned against using unadjusted mortality; analysis of the National Veterans Administration Surgical Quality Improvement Program data showed that classifying high and low outliers using unadjusted mortality could result in a 64% error rate. The same authors also noted that it has not been possible to carry out inter-rater reliability studies in the participating centres because of a lack of resources. However, they believed that the presence of the trained nurse auditors has ensured a high level of clinical credibility, as well as reliability and validity of data collection. The programme has the support of the staff involved, although some question whether administrative database analysis would not provide just as good information at lower cost.

An important argument for continuing with this special system is the need for the nurses to try to identify deaths outside hospital within the 30-day period of the definition.

These two systems are the main ones from which it might be possible to extract data on adverse events, although this has not been their prime intention. Other studies use the HMPS screening criteria for assessing quality of care. In 1998, Camacho and Rubin405 compared the validity and reliability of the HCFA’s Uniform Clinical Data Set System (UCDSS), the HCFA generic screens, and the HMPS screening criteria in a random sample of 451 patients undergoing percutaneous coronary transluminal angioplasty or CABG or experiencing an acute myocardial infarction at Johns Hopkins. To test reliability, two independent cardiologists applied reference criteria to the nurses’ screening decisions. The nurses screened the records using the three different methods. A further subsample was doubly assessed. Camacho and Rubin405 found that agreement between pairs of physicians on adverse factors was low ($\kappa < 0.40$). Between nurses, agreement on ascertainment was excellent for the UCDSS ($\kappa = 0.93$), slight for the HCFA screens ($\kappa = 0.11$) and fair for the HMPS criteria ($\kappa = 0.41$). Overall, the UCDSS was more sensitive but less specific, and took more nurse time (6.7 hours per quality problem, compared with 2 hours for the other screens). All systems were
better at detecting adverse events than quality problems and substandard care; the HMPS criteria were slightly more sensitive. Camacho and Rubin concluded that these systems perform better for the detection of clinical adverse events than for the detection of process or management errors, with the HMPS screening criteria and the HCFA generic scores performing the best.
Chapter 9

Discussion

This chapter incorporates and critiques the findings from the surgical mortality systems and generic, national monitoring systems (chapter 8). The quality of monitoring systems should be assessed in relation to their usefulness for the purpose for which they were designed. Many of the issues regarding the measurement and monitoring of clinical outcomes apply equally to surgical complications. A system for monitoring whether or not a rate is abnormal will require different content from one designed for individual feedback of single events. The former, for example, requires data from an unaffected but at-risk population that the latter does not. At an operational level, information should be complete, accurate and available at the right place at the right time. An early warning system has a different focus from one designed to allow comparisons with other centres or an analysis of possible causes. Since one of the main purposes here is to monitor change over time, information should be population based and time and space specific. The system should be parsimonious and collect only data for which a purpose is envisaged.

Lee and Pow surveyed expectations of clinicians for information and found that they wished information to be, in descending order of importance: accurate and precise; reliable; relevant; timely; and comprehensive. In respect of access to information, physicians put convenience before reliability, immediacy and responsiveness (i.e. ability to react to questions). Complete coverage is important to avoid the bias that will occur with voluntary incomplete participation; for example, in the original CEPOD study 13% of deaths were associated with the 5% of surgeons and anaesthetists who did not participate. Alderson identified a number of issues that are pertinent when using routine data rather than specially designed studies. These included: completeness; accuracy; delay in collection and processing; a coding system that is sufficiently specific for interpretation by clinicians; whether the content is appropriate for the purpose; sufficient flexibility to add items for local use; confidentiality; and, finally, an analysis and output that is acceptable to the potential user.

The assessment criteria for surgical mortality information systems were outlined in chapter 2. In summary, these consisted of: the use of standard definitions; coverage; denominator data; risk adjustment; data collection and validation; and output and feedback to clinicians. This final discussion is organised broadly according to these criteria.

Definition and ascertainment of surgical mortality

Several barriers to reliable monitoring of death as an adverse event have emerged from the review. First is the variation in the postoperative period covered by ascertainment of deaths. The definition and measurement of death in health-record systems may seem uncontroversial, and it is reasonable to think that it will be the outcome or adverse event most comprehensively recorded. However, it is clear that definitions vary in terms of the duration of postoperative period to be included and whether within that period deaths after discharge are actively sought. Unless the same period is used it is not possible to make valid comparisons of either incidence rates or the pattern of associated factors.

The most common definition is 30-day in-hospital mortality (Box 26). This definition is used by the specific audits of surgical mortality, the NCEPOD and the SASM, who both define death as up to 30 days, but there is no systematic attempt to ascertain deaths postdischarge. Only the SHFA includes deaths postdischarge, given as “death up to 120 days postfracture regardless of location” (Table 18).

Within the UK, cardiac deaths are monitored in three separate, but linked, systems that have three different definitions:

- the UK Cardiac Surgery Register defines death as “within 30 days of surgery”
- the UK Adult Cardiac Surgery Database defines death as “on the same admission as surgery in the base hospital”, regardless of postoperative timing, so that some may be later than 30 days
Discussion

• the UK Heart Valve Registry, which is linked to the national death registries of all four countries, captures death at any time, and can therefore ascertain a true 30-day mortality.

In the USA, the largest cardiac database, the STS Database, defines death as “all deaths occurring during the hospitalisation in which the operation was performed”. It also attempts to include deaths after discharge if they are clearly related to the surgical procedure, but not in any systematic way. Valvular operative mortality is defined as “within 30 days”, but has a separate category of late mortality. The Veterans Affairs Continuous Improvement in Cardiac Surgery programme also defines mortality as “all postoperative deaths within 30 days, in or out of hospital, plus any later than 30 days as a direct result of surgery”. Finally, the SWEDVASC defines mortality as death up to 1 year postoperatively, but the registry is linked to the national death registry and is, therefore, informed of deaths at any point after surgery, which could be years.
In routine or administrative information systems, the period of death ascertainment depends on the linkages that exist between different parts of the healthcare network. In the UK there is no linkage between hospital and primary care, but it is possible through the national Registrars of Deaths to request flagging and automatic notification of relevant deaths and this is routine in Scotland. In 1993, Smith linked hospital discharge data with the Registrars' vital statistics system and found that 32% of perioperative deaths occurred after transfer to another specialty or hospital, 8% on re-admission and 8% outside hospital.

There are different approaches to identifying heart-valve deaths: first, they can be identified from the heart valve register, which captures death at any time; and, secondly, they can be identified from the death certificate. Morton and co-workers found that the presence of heart valves was recorded on the death certificate of only 21% of patients with a valve prosthesis, which suggests that flagging of cases is the only way in which ascertainment can be accurate. Case ascertainment is likely to be far higher when determined from the procedure-specific database.

Within the USA, one of the main reasons for lack of follow-up is the enormous difficulty of tracking patients and the legal barrier to a common identifier. The position is similar in the UK, but is surmountable by record linkage with death records. In US organisations that provide care in different locations, there is the potential to link information. For example, the HCFA is now being linked to the National Death Index, although this index is reported to be slow, expensive to use and does not provide information on the cause of death. As a consequence, Newman and Brown evaluated the ability
of a microcomputer program, Automatch, to link patient records in a hospital database to Californian state mortality files and the US Social Security Administration. Using hospital information about known deaths, they found the Californian death tapes to have a sensitivity of 99% when the social security number was available, compared with 86.6% when it was missing. Sensitivity, therefore, was nearly perfect. Secondly, they evaluated the death registries as a means of ascertaining mortality, and found that sensitivity was less for deaths outside the index hospital (81–86%), but that they could approximate the mortality rate in patient groups with a high risk of death. No information was found on whether this linkage system has been pursued.

In a Belgian study, extensive follow-up procedures after coronary surgery were undertaken in 1987 and 1993 to compensate for the lack of comprehensive longitudinal records. Patients thought to be alive at a common closing date were pursued by post, telephone, or through neighbours or the family physician. If patients had moved house, the town hall was contacted for their new address. If the patient had died, attempts were made to speak to someone who was present at the death, to define the mode of death. The second of these studies found 99.9% of 9600 patients, but it took more than 2 years from the mail date of the first letter. The authors commented that this costly and time-consuming process to find postdischarge deaths contrasted sharply with the present intensive scrutiny of in-hospital outcomes. Moreover, compared to a widely assumed 30-day post-CABG mortality of 2% as the total operative mortality, they demonstrated a continuing risk of attributable death up to at least 2 months postoperatively. There was also a 100-fold spread in the 30-day mortality, but most of the mortality was concentrated in a few patients with severe pathology and co-morbidity. Thus, the authors concluded that risk adjustment is crucial to interpret avoidability.

One German study examined early mortality after CABG over a 10-year period and found that a substantial number of deaths occurred between 30 and 60 days postoperatively. When data were divided into periods (1988–1991 and 1992–1997) they observed that 30 days was sufficient to detect most of the deaths in the earlier period, but that this was not so in the second period. A multivariate analysis of preoperative concomitant factors identified that there were more patients in the second period who were at higher risk because of co-morbidity and poor preoperative status, yet much of this did not emerge until after 30 days. They observed that this type of shift should be expected as procedures become established, and that there is a need for both postdischarge follow-up and continuous adjustment for changing case mix. Based on this analysis, the authors recommended that follow-up after CABG should be extended to a period of 180 days.

Interpretation of when it can be attributed is also problematic. To address the question of the duration of ascertainment of deaths that might reasonably be attributed to the operation and not to background risk factors, Seagroatt and Goldacre used different methods to calculate and compare postoperative fatality rates. These methods included calculation of in-hospital mortality, standardised mortality ratios (using the resident population of the hospital catchment area), procedure-specific postoperative deaths (case fatality rates) and the ratio of early (<30 days) to late (90–364 days) deaths (the ‘relative mortality ratio’). In-hospital and case fatality rates within 30 days were similar for emergency admissions, and rates were approximately double those for elective procedures. Similarly, the temporal comparison of early to late deaths demonstrated the relative mortality ratio to be very different between elective and emergency procedures, and that for most elective procedures there was no excess mortality after 90 days. The authors suggested that 30 days was a reasonable cut-off for attributable postoperative mortality. However, they argued that the use of procedure-specific temporal profiles and relative mortality ratios allows operations to be grouped into those with and those without associated mortality, provided that emergency and elective operations are analysed separately, suggesting that these analytical methods might be of use in surgical audit and confidential enquiries. Although this model was appropriate for the surgical procedures included in the study (e.g. prostatectomy, herniorrhaphy, cataract, hip replacement) it would not be appropriate for operations where fatality rates do not stabilise over time (e.g. cardiac surgery, cancer surgery). No further uses of this potentially useful approach were found. The proposed relative mortality ratio deserves further testing in different populations.

It is clear that 30 days is the most common, but by no means universal, arbitrary cut-off point for case ascertainment of surgical mortality. Very few systems have a complete follow-up procedure and
only those linked to national death registries can guarantee reasonably accurate ascertainment.

**Risk adjustment**

Risk adjustment is a fundamental feature of the US monitoring systems reviewed here, particularly the cardiothoracic registers. The original purpose was for accurate prediction of each hospital's case-fatality risk so that any higher than average rates could be justified and blame for poor surgery was not attached to either a hospital or an individual surgeons. The position in Europe was much less defensive, at least initially, but it is clear from the literature that risk prediction or adjustment is now becoming the norm. Risk models have become very sophisticated and, although this was not a part of this review, several widely used models were identified, including the Parsonnet risk score, POSSUM, EuroSCORE and APACHE. Risk varies considerably by age, severity and co-morbidity. If the system is to produce rates instead of, or as well as, simply identifying individual events for discussion, it is essential that the rates be risk specific. Without risk adjustment, the epidemiological information that UK audit systems, such as the NCEPOD, have produced is much less useful than it would otherwise be in identifying where and for whom changes in care are needed. As found within infection monitoring systems, risk adjustment highlights those patients who may require special monitoring or intervention. Moreover, participation of surgeons and anaesthetists, even if it becomes compulsory, will be more productive if they can identify with the data that are produced. That will be much more likely if they can see the effects of the risk factors that they know clinically. The review group, therefore, recommends that risk adjustment be part of any system for monitoring surgical adverse event, particularly mortality, and that a separate review be done of the different risk models that exist.

**Denominator data**

The ability to relate numbers of surgical deaths or adverse events to a population at risk is crucial if what is of interest is the incidence rate in that population. There are several relevant populations of interest. First, and most useful, would be the numbers of individuals put at risk by having the surgical procedure (e.g. the number of relevant operations that were done). This allows the calculation of case-fatality rates in a defined exposed cohort. Second, is the number of individuals reported to the monitoring system. This will be biased by the extent to which coverage is incomplete and the patients included do not represent the totality of those who had the operation. Third is the population of the community served. This figure is probably of less interest to surgeons, and the resultant rates are influenced by stages of care before surgery, but it is useful for epidemiology and planning to assess the impact of surgical iatrogenic disease on that community. Ideally, if the purpose of the monitoring is to identify where and how care should be improved, a denominator should also allow description of non-operated (or, in the case of death, surviving) individuals in sufficient depth to know how they differ from those operated on in respect of risk. This allows an estimation of relative risk of death as an outcome of surgery.

All systems for cardiovascular surgery collected denominator data on procedures, and thus reliability depends on the extent of coverage. Only the SWEDVASC appeared to be close to complete coverage and could consider population levels of surgical mortality or surgical adverse events. Incomplete voluntary systems permit calculation of risk, but no estimate can be made of bias due to non-participation. Generic organisational systems, such as the Veterans Affairs, HCFA and those in the UK, contain the denominator of all recorded relevant operations to act as a background rate of exposure. This background rate was seldom used, except by the NCEPOD in their first report. Systems with national or state coverage can provide an approximation of the third calculation, population relative risk. Such systems include SWEDVASC, the British Columbia Provincial Cardiac Registry and the systems used in the four countries of the UK.

Thus the ability to relate events accurately to populations at risk was limited. Time trends of rates, therefore, mean very little because they depend on the completeness of reporting of operations from year to year. Arguably it might be preferable not even to attempt to produce rates, but rather to do as, for example, the NCEPOD, the SASM and the UK Cardiac Surgical Register have done, and restrict reporting to descriptions of factors associated with the deaths that were reported, leaving interpretation of the relative occurrence and meaning of these factors to local review. This approach may be especially important when dealing with small numbers of surgical deaths or complications.
**Morbidity**

Surgical morbidity *per se* was not included in this review of mortality monitoring systems, although a vast range of audits and morbidity systems were found during the search process. The surgical systems that covered morbidity as a complication included SWEDVASC and the STS Cardiac Surgery Database. The other systems recorded morbidity either for the calculation of risk or as another item in a generic system. However, the generic systems do not usually distinguish between preoperative risk morbidity and postoperative complications. Silber and co-workers found that using complications to rank hospital performance gave different and less reliable comparisons because of the poor quality of the morbidity data. The assessment of the quality of postoperative morbidity data was reiterated by Grover and co-workers in a critique of the Veterans Affairs system. It is a recommendation that data collection distinguishes clearly between pre- and postoperative morbidity.

**Data processing and validation**

It is difficult from the material used in this review to present the characteristics of an ideal data capture and processing system, as only a few articles were found that focused on the system itself and thus comparison of the effects of different data procedures was not possible. However, a number of evaluations were found of components of a data processing system, and in particular of the accuracy and completeness (i.e. reliability) of data capture and entry.

It is fundamental to have complete recording of the required clinical information in the primary record. Over the years, the Scottish hospital records system has carried out numerous studies of data completeness and accuracy. The consistent conclusion is that, until clinicians take the time and effort to complete the clinical information, no amount of support staff or changes in procedures will improve the quality of information for audit and research. Ironically, clinicians often berate the quality of the data produced by the systems. Although much desired, clinical information, especially on risk factors, is usually seen to be the least well completed, although in Scotland the gap was mainly in secondary diagnoses and secondary procedures. Ascertainment in FINNVASC was worst for emergencies and endovascular procedures.

Several systems noted that having dedicated data managers rather than doctors, whether or not they were nurses, improved the completeness and accuracy of records. In the USA, data managers are seen as a check on gaming. Many systems believe that using dedicated local reporters allows special training and time for chasing events in the wide range of records that may have to be searched. This may be especially true when data are to be extracted from routine records. Timing of data collection has received few comments. There is evidence that coding immediately after the death is more sensitive and specific than waiting until after the discharge letter has been written. Similarly, it has been shown that the most accurate data capture was at the time of the operation. Valedictory checks built into the data processing software are the norm for internal system validation, by checking for logical errors (such as ‘pregnant male’), as is central–local completion (UK Heart Valve Register), and an interactive website for data managers appears to be much appreciated by those in systems that have one (e.g. STS databases). Volk and co-workers found that reliability was improved by providing definitions, if an actual number or clinical value was requested, and by asking for precise values rather than ranges.

One of the four new quality control measures introduced in 1996 by the Audit and Validation Committee of the STS was a national nurse coordinator to provide educational activities, and the Scottish systems use quality control coordinators for training and to ensure consistency across centres (SASM, STAG). No reports were found of the impact of these measures. SWEDVASC, the Scottish system and most US systems reported reproducibility checks by re-registering a sample of records, or by an external review or site visit. Inter-rater reliability is evaluated periodically in several systems. It is not clear whether automated data transfer improves the reliability of the data capture, although, intuitively, if it reduces one stage of extraction and transcription it seems that it would improve capture. Several systems, especially in the USA, enter source registries in total, but no validation has been published.

Thus, while it is possible to assume that certain processes will improve the reliability of data capture, published evidence is in short supply, and most dedicated UK systems state that they are lacking external validation studies. From the published literature there is an impression that
SWEDVASC is close to a model of validated processes, but it may be that some of the US cardiothoracic systems are equally well validated but have not reported it to the same extent. Bergqvist and co-workers\(^3\) noted several factors that influence the quality of data:

- There are numerous opportunities for the human factor to influence ascertainment and none of the registries was a gold standard. Local surveys are needed to improve completeness.
- The most frequent reason for failure to report on follow-up is that patients are taken care of by other physicians, but other reasons are difficulty for the patients to get to the hospital and the varying interest of surgeons. SWEDVASC sends monthly lists of patients to hospitals, in part as reminders.
- Half of the hospitals now have their own computer-based routines from which information is electronically forwarded; these systems can compel the surgeons to enter all data before the whole entry is accepted. However, this can be avoided by entering ‘don’t know’.
- Most incorrect information concerns risk factors.
- Changes to the protocol should be kept to a minimum, both for consistency and to permit comparisons over time.

It appears that dedicated local data collection staff improve both completeness and accuracy, but that they need to have better methods of prompting clinical staff to provide the necessary data. We recommend that this specific aspect of data collection be examined more closely and new methods evaluated.

**Costs**

It is difficult to compare costs between systems because the denominators vary, as do the calendar years of the reported costs. However, none of the systems reviewed here appears to be excessive in terms of cost in comparison to the perceived value of the information in improving care and, ultimately, reducing deaths. The SASM is quoted at around £40 per death and SWEDVASC at around US $6 per case. The STS in the USA quotes a figure of US $30 per chart entered, and the same figure applies to the Northern New England Consortium. Within the Veterans Affairs system, costs are quoted as US $12 per procedure assessed and $38 per major procedure assessed. Nugent\(^4\) observed that US $30 per chart entered was less than 1% of the total cost of one CABG. He also calculated that local costs were about 50 cents per chart.

**Output and feedback**

With regard to output and feedback, the issues relate to: the level of detail (i.e. whether by hospital, unit or clinician); delays in turnaround; public accessibility; and impact of monitoring systems. While most systems collect data on surgeons and anaesthetists, none publish an analysis by individual clinician, although several give individual feedback if it is requested by the person concerned. This emphasises the self-auditing purpose of most systems. The HCFA, who originally published by both hospital and cardiothoracic surgeon, no longer do so. Dedicated procedure-specific feedback was preferred\(^5\) and was found to be more user-friendly than generic information that needed further interpretation.

Even in efficient systems delays in producing analyses of aggregated data from a central point arise because the pace is set by the slowest participant, and validation requires data interchange with contributors to correct obvious errors. It was only possible to judge delays in publishing reports, either paper or website, and the average appeared to be about a year at best. Local feedback is of course faster, and early unadjusted feedback was reported to be popular. However, there was little other than anecdotal comment about the perceived best quality and user-friendliness of output, and no obvious criteria by which this should be judged. Public accessibility to data has not been an issue, apart from the wave of concern, already described, over the HCFA’s original publications in the late 1980s.

The impact of monitoring is even more difficult to assess. Only a few of the publications reviewed in this report described the impact of their existence. Monitoring has been shown to lead to avoidance of high-risk patients. Schneider and Epstein\(^6\) surveyed a group of Pennsylvanian cardiologists and cardiac surgeons about their perceptions of the usefulness of a consumer guide to CABG. Although few of the respondents rated the document as very important in assessing surgeon performance, 63% of the latter reported that they were less willing to operate on severely ill patients. A similar survey in New York found fewer cardiologists who had changed their referral practice and did not think that avoidance of high-risk
patients could account for the falling CABG mortality rates in the USA.421

Epidemiological systems that analyse for risk factors, such as the NCEPOD and the SASM, reported changes in clinical practice and facilities that reflect their recommendations, but it is extremely difficult to make a direct cause-and-effect link. Reports that influence individuals to improve their own practice are unlikely to be visible in the data set unless they were extreme outliers or they publish it, as was done by de Leval and co-workers.422 Bergqvist and co-workers,393 reporting on what appears to be one of the best evaluated systems, said that it was not possible to assess the influence of SWEDVASC on improving care. In response, Ruckley440 observed the difficulty of assessing the influence of the registry on the quality of Swedish vascular surgery. He stated “in reality, an influence on the quality of surgery as a result of this type of registry will generally be a gradual process and difficult to distinguish from a range of other influences such as the published literature, professional intercommunication, technological development and natural evolution of the specialty”. Thus it is not possible from the literature to make recommendations about the most effective or liked form and content of feedback, or of its impact on surgical care. A separate review of literature related to ‘quality improvement’ and subsequent impact on care, however, might yield evidence not found in the surgical monitoring literature.

Generic versus dedicated monitoring systems

There has been considerable debate in the literature about the relative merits of using existing institutional generic monitoring systems to extract the relevant items on surgical adverse events versus a dedicated condition-specific system for the procedure or complication of interest. The option of generic information systems rests on the facts that:

- they do not require the additional work of a dedicated auditing or monitoring system
- they are generated as a by-product of routine care and are therefore relatively inexpensive
- they have the potential to detect adverse events that present outside surgical services423
- they are usually large enough to reduce the likelihood of selection bias
- they span a period of time, thus allowing trend calculations
- they have diverse items in them that may allow the search for unexpected associations.424

In the USA, Peterson and co-workers441 used Medicare (HCFA) claims data to assess secular changes in CABG and percutaneous transluminal coronary angioplasty. They considered that claims data had three main strengths: size, inclusiveness and continuity of collection. The size of claims data allowed analysis of variation in outcomes by minority groups, who had previously been underrepresented in publications. Lillard and Farmer425 also discussed the research potential of the very large HCFA database linked to a range of national survey data, and especially the National Death Index.

The question at issue is whether generic systems meet the criteria for a good monitoring system sufficiently well to be useful and useable in the detection of surgical mortality and surgical adverse events. Hannan and co-workers416 summarised the disadvantages of what they termed ‘administrative data’:

- the inaccuracies of the ICD-9 CM codes used for diagnosis
- the inability to determine the timing of an event (i.e. whether it was a co-morbidity present on admission or developed postoperatively)
- the absence of important information such as clinical risk factors, functional status and quality of life assessment
- some more complex diagnoses are coded than are needed for reimbursement
- the number of diagnoses is limited to five, which may not be enough to interpret complicated cases.

To reduce these drawbacks, they recommended consideration of a prospective clinical database (i.e. the system now used by the STS and others), but acknowledged that such systems are expensive to develop and maintain. They also cautioned that their assessment of the balance of advantage of a clinical database over an administrative system for monitoring cardiac surgery performance would not necessarily apply to other conditions and types of patients, presumably because the need for risk adjustment might be less.

A related dimension of information systems is whether they are local or part of a central network that shares data to provide comparative information. Shaw408 argued that local studies usually have advantages over national ones because there is better control and the quality of the data is
higher in terms of criteria such as accuracy, completeness and confidentiality. However, for most death monitoring the number of events is too small to permit statistical interpretation. Audet and Scott\textsuperscript{426} noted that national data sets such as the HCFA Uniform Clinical Data Set were useful for national descriptions of practice patterns, but that detailed assessment of quality was better done at local level.

Only one local system was found. Aberg and co-workers\textsuperscript{427} described a dedicated single-hospital system for registering "deviations from the norm" in cardiac surgery that appears to meet all the criteria for a well-designed system. It avoids duplication and redundancy of recording, is used for patient administration also, has measurable definitions of postoperative morbidity with defined levels of severity, and has been demonstrably and rapidly effective, through monthly feedback, in reducing the incidence of thrombophlebitis. Unfortunately, no published evaluations were found of the completeness and accuracy of data collection.

Overall, fewer than ten studies were found that assessed the relative performance of the different approaches to monitoring. The absence of clinical items for risk adjustment in administrative data was a common theme. Krakauer and co-workers\textsuperscript{428} compared the HCFA's claims data with a clinical risk adjustment model using the same data at source plus selected clinical factors. They found that the clinical model was a better predictor of the probability of patient death. Using the clinical model as the reference standard for the prediction of outlier hospitals, they found that the claims model had a sensitivity of 81\% and a specificity of 79\%. However, its positive predictive value was low (i.e. it had a high rate of false-positive results). In respect of hospital ranking, the two models performed similarly. Krakauer and co-workers\textsuperscript{428} concluded that claims data were useful in describing the distribution of mortality rates across hospitals, but less so for pinpointing an outlier.

Similar conclusions about the need for clinical data to improve prediction were reached by Audet and Scott,\textsuperscript{426} Hartz and co-workers\textsuperscript{429} and Hannan and co-workers.\textsuperscript{416} Some UK surgical audits were developed because of a belief that the national generic systems were not fit for their purpose because of inaccuracies, inadequate operation codes and lateness of production.\textsuperscript{351,358} The SASM attributes its success to the special operation coding system that it has devised,\textsuperscript{361} while the Scottish Arthroplasty Register has judged that it is easier, and is sufficient, for it to use the national system.

In 1994, Cleary and co-workers\textsuperscript{442} conducted a study whereby routinely collected data from the hospital information systems of five London hospitals were compared with data retrieved from case note review. Trained abstractors extracted at least double the diagnostic codes retrieved from the hospital information systems. They also noted that records abstraction by an experienced nurse was sufficient to identify adverse events, although this was not feasible using the hospital database. Thirdly, he found that the absence of secondary diagnoses (i.e. co-morbidities and complications) from the database was not wholly due to the failure of clinicians to record them, because they were there but not extracted.

Crucially, however, attempts to encourage the use of the Scottish national system with add-on items for specific clinical audits have encountered what appears to be the main barrier to using generic systems; namely the failure of clinicians to complete the necessary clinical items in either the case record or the discharge summary on which the national system is based.\textsuperscript{418} Thus it is difficult not to conclude that motivation may be more important than any technical attributes of the different systems. Whether there is potential to tackle this through surgical accreditation is a matter for discussion and debate, or whether the answer is an investment in dedicated data managers.

Is a combination of systems possible? Spiegelhalter\textsuperscript{2} argued that it is both possible and beneficial to integrate epidemiological and clinical approaches to monitoring surgical adverse events (i.e. monitoring of groups or rates versus identification and analysis of individual case histories). He referred to the HCFA's quality improvement initiative, which is an amalgam of the early hospital-specific mortality rates and peer review of individual case records. An example quoted by several authors is that of de Leval and co-workers,\textsuperscript{422} who found a worryingly high number of deaths in their own series on neonatal correction of transposition of the great arteries, and developed a sophisticated analysis of their own data to identify where improvement had to occur. Spiegelhalter used de Leval's work as an example of both clinical and epidemiological approaches in the great tradition of Codman and Nightingale, while Lovegrove and co-workers\textsuperscript{430} used it as the basis of devising a relatively simple risk model that takes account...
of both case mix and *a priori* risk information. Often recommended is a combination that identifies deaths or other surgical adverse events locally for individual discussion before submitting data for national collation; we did not find any published examples of this, although anecdotal evidence suggests that it does occur. For both avoidability and morbidity complications that occur in small numbers, such a combination might permit both qualitative and quantitative assessment within the same system. It might also permit early warning of unexpected events. We recommend that this combination should be tested as a matter of urgency. A separate 3-year HTA methodological review assessing ‘critical incidents’ within medicine is currently underway.

**Summary**

Ultimately, the choice of a monitoring system depends on purpose. The advantages and disadvantages set out above continue to be real, and focus on two main ones: clinical content and timing or speed of feedback. If the aim is to create statistical risk models that identify outliers or risk factors, then ideally the quality of data should permit exception (e.g. reporting when a case or rate falls outside the expected pattern of care), and accurate clinical detail is essential to a level that few generic or organisational systems could justify for all conditions. If, on the other hand, the aim is to inform a peer assessment process based on local discussion against comparative information from other sites, then extraction from a generic information system is an efficient screening process. However that is all that it can be; a screening process to identify those patients who probably have an adverse event, or factors that are probably associated with a higher than average risk of death following a particular procedure. In almost all instances this will require a case-based follow-up to confirm and explain or attribute a surgical adverse event.

Certainly in the UK, fitness for purpose will increasingly be judged by the ability to detect unexpectedly high levels of adverse events at an early point. The inevitable delay in central analysis means that no system can act as an early warning system unless it is adapted at a local level to pick up something unusual before submission to the central point. It is likely that such monitoring will be of single events rather than of incidence rates. Moreover, there can be no automatic assumption that even if the death, in terms of mortality monitoring, is attributable it is also avoidable in any one individual and therefore a true adverse event. If audit and sanctions are local then the need for statistical risk adjustment models is less. However, clinical governance suggests that local sanctions are not going to be enough. Therefore either the monitoring system will have to be very much more sophisticated and expensive than is currently the position in the UK, or avoidability will be defined by discussion, not calculation.
Chapter 10

Conclusions

This methodological review set out to establish valid and reliable definitions of surgical adverse events and to explore systems for the measurement and monitoring of unwanted consequences of surgery. Surgical events were selected on their frequency of occurrence in national hospital and surgical audit data and for the likelihood of evidence of measurement and systematic monitoring. A cross-section of common and rare postoperative events was included. Surgical wound infection is a frequently occurring complication that seldom leads to mortality; conversely, anastomotic leak is a less frequent event, but it is associated with significant morbidity and mortality. DVT is also associated with significant postoperative morbidity and mortality. Surgical mortality was included as an irrefutable, quantifiable outcome and was used in conjunction with surgical wound infection to identify examples of surgical monitoring systems. Table 19 presents an overview of the review findings for each of the four surgical adverse events.

Attempts have been made over the last decade to establish valid, reproducible and repeatable definitions of surgical wound infection. The development and evaluation of wound classification systems, particularly within the UK, have added to the theoretical body of knowledge. In relation to the monitoring of surgical wound infection, hospital-based systems in the UK have adopted US definitions, protocols and standard methodologies for data collection, processing and output. However, the accurate detection and measurement of surgical wound infection after discharge remains a challenge, particularly as many hospital-based monitoring programmes exclude patients undergoing day-case and short-stay surgery. No single valid, reproducible definition or measurement of anastomotic leak was identified from our literature search. Diagnosis of anastomotic leak is based on a combination of clinical judgement and radiological assessment. The definitions and values used to measure anastomotic leak vary extensively, such that comparison of rates between different studies and institutions is difficult. More rigour is required, particularly within multicentre trials, and collaborative attempts should be made to develop standard and accurate definitions.

Although mortality is consistent in the definition of the event itself, there is variation in the duration and follow-up of postsurgical case ascertainment, such that comparability of mortality between systems is difficult. In-hospital mortality rates differ in whether they cover in-hospital transfers and late deaths. Thus each topic reviewed demonstrated variability of definition and measurement, such that comparisons across time or location would be unreliable.

At all levels, external evaluation of the accuracy and completeness of capturing data from clinical records, even in hospital, is in short supply.

A number of issues arose from this review that can be generalised to the measurement and monitoring of surgical complications as a whole. Clearly, the epidemiology and time frame of development of individual events needs to be understood. Certain events, such as death, are more likely to occur in the hospital setting, and therefore postdischarge surveillance can be undertaken as an occasional or intermittent audit. Other events, such as surgical wound infection, are more likely to develop after discharge, and thus manifest in the community setting where monitoring and surveillance proves challenging. There is a move towards procedure-specific reporting, although time limits should be agreed for individual events (e.g. 30-day surgical wound infection rate). Ideally, these limits should not be arbitrary, but rather should be based on an understanding of the natural history of the complication, including its chronology. However, in the absence of this knowledge, agreement on an arbitrary time span of ascertainment is better than arbitrary variation. Relevant preintervention risk factors should be recorded and distinguished from postoperative complications. This approach has been adopted by established, dedicated monitoring systems, such as the cardiothoracic databases and nosocomial infection programmes.

Linkage to existing monitoring systems could be undertaken at different levels. First, where the event is death, linkage could be achieved between hospital surgical systems and national death registers, provided that the linkage can be done acceptably within data protection legislation.
However, the quality of detail on death certificates will guarantee accurate information only on the event of death rather than on its cause and any secondary diagnoses. Second, where an event is severe enough to require re-admission for further treatment or management, including re-operation, links within existing systems should be pursued. There are examples within national routine systems of linkage between the primary admission for surgery and subsequent re-admission with an associated adverse event. Further work could be conducted to develop this method for tracing late events and unwanted consequences of surgery. Again, the period of time between hospital re-admission and event occurrence will depend on the epidemiology, natural history and timing of individual adverse events. Thirdly, there are events that manifest and resolve within the community setting, either with or without management from primary care staff. Linkage between primary and secondary care information systems is rare and currently occurs only when the same organisation manages care in both settings. However, the moves within Europe to establish a single patient identifier may make such linkage easier, again provided that it can be done within the terms of data protection and human rights legislation. The challenge here is balancing the resources, in terms of effort and cost, for accurate monitoring.

Finally, in the systems reviewed here, and these are the main surgical audits or information systems identified through the Royal Colleges, we found little evidence on the benefits of monitoring in terms of its impact in improving care. However, there is a hint that rapid and local feedback is more likely to be followed up to identify reasons for unexpected levels of complications, and any required changes are more likely to follow. Thus it is possible that a combination of local immediate feedback of events to local clinicians and subsequent central aggregated calculations of rates, risk and trends would be the most useful and acceptable approach. There are no examples where attribution is apportioned to antecedent surgery, unless assessment is undertaken at the level of the event or individual. The question is which one is more likely to lead to improvements in surgical care and outcomes?

In conclusion, at a strategic level the trade-off between benefits and costs is fundamentally important to the choice of approach to be taken in monitoring adverse events as a whole. In the past year, there has been an explosion of public and professional interest in ‘medical errors’ and it is important to understand the difference between the two approaches. The generic definition for adverse events used by many of the medical error advocates is “an untoward event which results in prolongation of length of stay or disability at the time of discharge” (e.g. HMPS\textsuperscript{431}). In contrast, the epidemiological approach advocated in this review covers all events, regardless of their severity or their impact on patients and their care. Medical error, as presented in the literature to date, focuses on identifying and acting on the reasons for an unexpected harmful event to a single patient, and is only secondarily concerned with examining the rates of such events in an epidemiological or statistical framework. It takes an organisational approach, as set out comprehensively in An Organisation

| TABLE 19 Overview of the measurement and monitoring of selected surgical adverse events |
|----------------------------------|---------------------------------|-----------------|-----------------|-----------------|
| Review criteria  | Surgical adverse event | Surgical wound infection | Anastomotic leak | DVT | Surgical mortality |
| Definition        | Clinical/bacteriology | Clinical/radiological tests | Clinical/diagnostic tests | Clinical | Timing variable |
| Measurement (subjective/objective) | National hospital statistics, nosocomial systems, surgical audit\textsuperscript{a} | National hospital statistics, surgical audit\textsuperscript{a} | National hospital statistics, surgical audit\textsuperscript{a} | National hospital statistics, SASM–NCEPOD, specialist registers |
| Monitoring: in-hospital | Postdischarge | None | None | None | Linkage to death register\textsuperscript{b} |
|                      |                   |  |     |     |                      |

\textsuperscript{a} Intermittent local/regional audit rather than continuous national monitoring

\textsuperscript{b} Not universal
with a Memory\cite{432} rather than a clinical one and, indeed, is directed at moving attention away from the clinician as the source of error. To date it has dealt mainly with possible errors in drug administration and failures of the system to deliver the right care at the right time and place. Our epidemiological approach means that the incidence of events will be more likely to be true, higher than those detected only by self-reporting rather than monitoring, and capable of quantification as a rate. These aggregated rates draw attention to outliers, whether they be at the individual, unit, department or institution level, but within a statistical rather than observational framework.

However, although current emphasis is on acting to reduce ‘one-off’ errors in health organisations, ongoing monitoring of adverse events and complications at the clinical core of the service remains in dire need of attention if it is to underpin both clinical improvement and avoidance of error. All the advocates of a medical error approach concur that agreement does not exist across healthcare units on the definition of what is an adverse event, whether a clinical complication or a systems error, and that without this the search for remediable causes will be hampered considerably. On the one hand is accurate and continuing information, fed back primarily to the clinical process as part of audit and clinical governance to allow improved clinical procedures. On the other is volunteered information about overt critical incidents or near-misses as identified by any individuals in the healthcare organisation. It is relevant that very few of the published critical incidents have been surgical; this may be due in part to the fact that although a complication is regarded by clinicians as important it may not be highly visible within the system of care. The trade-off is between focusing on organisational mishaps or seeking continuous clinical improvement. Arguably, both are needed, but it is important to understand that they are very different.

Thus this review, which is a move towards bringing the monitoring of surgical procedures into line with that for medicines and devices, is a small but important step towards a more rational identification of surgical complications in a rapidly changing clinical field.
Chapter 11
Implications and recommendations

The review group has highlighted areas for future work based on the gaps in the published literature. These recommendations are presented according to the four specific review topics.

Surgical wound infection

- There is a clear need for surgeons and other healthcare professionals to adopt a single, standard definition of surgical wound infection if comparisons over time and between departments and institutions are to be valid, accurate and useful. The use of the 1992 CDC definition for superficial incisional, deep incisional and organ/space surgical site infection should be considered for use in the UK by hospital monitoring programmes and surgical audits. **Rationale:** This systematic review identified 41 different definitions of surgical wound infection from literature published over a 7-year period.

- Research is needed on the performance of the CDC definition in the UK setting. **Rationale:** Given that there are gaps in the body of knowledge in relation to the performance of the 1992 CDC definitions for superficial incisional, deep incisional and organ/space surgical site infection in the UK setting, consideration should be given to differences in surgical wound infection rates according to the CDC component used; in particular “diagnosis by surgeon or attending physician”.

- There is a need to assess formally the reliability of self-diagnosis by patients. **Rationale:** Little work has been done in the UK on the ability of patients to recognise signs and symptoms of infection. Given the need to set up post-discharge monitoring, this is clearly an option to be explored.

- There is a need to assess formally the reliability of case-ascertainment by infection control staff. **Rationale:** There is no published literature on the reliability of case ascertainment by infection control personnel in the UK and there are no published comparisons between infection control and surgical staff. There is evidence from the USA that dedicated staff are more effective and efficient at ascertainment.

Anastomotic leak

- Work is needed to create and agree a standard, valid and reliable definition which is acceptable to surgeons. **Rationale:** Leak is judged by surgeons to be the major complication of bowel surgery and it carries a risk of death. Reliable evidence of its associations in different clinical settings and units is needed before it can be prevented or at least its impact greatly reduced.

Deep vein thrombosis

- A separate systematic review is needed of the different diagnostic tests for the diagnosis of DVT. **Rationale:** Although the critical review of the DVT literature was not completed within the realms of this review, it was evident that a number of new techniques for the detection and diagnosis of DVT have emerged in the past 20 years. Our recommendation is that a separate systematic review be undertaken of the different diagnostic tests.

- The following variables should be considered in any future DVT review: anatomical region (lower limb/upper limb/pelvis); patient presentation (symptomatic/asymptomatic); outcome of diagnostic test (successfully completed/inconclusive/technically inadequate/negative); length of follow-up; cost of test; whether or not serial screening was conducted; and recording of laboratory cut-off values for fibrinogen equivalent units.

Monitoring systems

- A critical review of surgical risk scoring systems is urgently needed. **Rationale:** There was not time to include it in this review, but there is a lot of literature and an impression that surgical groups are devising new risk scores before evaluating existing ones. This has implications for the identification of truly excess events and attribution of cause and effect.

- In the absence of automated linkage there is a need to explore the benefits and costs of monitoring in primary care. **Rationale:** There is some evidence on the benefits of alternative...
methods of postdischarge surveillance, but this needs to be put in the context of UK primary care and the costs for primary care staff are not clear. However, for certain conditions, such as wound infection, the majority of events occur in this setting and are not being ascertained.

- The growing potential for automated linkage of data from different sources needs to be explored as a means of improving ascertainment of surgical complications, including death. It should include linkage to primary care, the private sector and death registers. However, this has to happen within the terms of data protection and human rights legislation. **Rationale:** (a) Shorter lengths of hospital stay and the expansion in day surgery mean that more surgical complications are occurring after discharge; (b) private–public partnerships are likely to push more surgery to the private sector (at present data on surgical procedures in private hospitals is obtained rarely and only voluntarily); and (c) routine linkage to national death registers would permit charting of the chronology of postoperative deaths. If linked to clinical data, there is the potential to develop procedure-specific case-fatality rates. However, a crucial accompanying step would be the availability of aggregated denominator data of the population at risk (i.e. all who had undergone the procedure of interest). Moreover, linkage is only possible if it complies with the need to use personally identifiable data only for linkage purposes and within a secure system.

- A review is needed of the extent of use and efficiency of routine hospital data versus special collections or voluntary reporting. **Rationale:** This review did not examine the extent of data collection for audit and other monitoring in hospitals. However, the recent shift in focus of the literature on adverse events, the renewal of clinical and corporate governance, and repeated comments in the publications reviewed here all suggest that the purposes and efficiency of hospital data collection require reassessment. Specific items that merit review include: (a) the role of clinicians compared with that of dedicated data collection personnel; (b) the combination of local early warning of an event and background central monitoring using verified data; (c) the use of risk scores; (d) whether special collections for specified topics and time periods would be more effective than continuous monitoring; and (e) whether mandatory participation in national audit schemes would yield more demonstrated impact on clinical practice.
Acknowledgements

This study was commissioned by the NHS R&D HTA Programme. The authors are indebted to the HTA referees for their perseverance in reading this report and the quality of their comments. The views expressed in this report are those of the authors, who are responsible for any errors. The authors would like to thank Mr Ghassan Al-Khairy, Mr Garth Reid and Ms Emma Pitchforth for assistance with retrieval and photocopying of articles, Mr Kenneth Park and Mr Dean Phillips for input at an early stage in the review, and Ms Susan Campbell for assistance with critical appraisal. We are also indebted to Mr Steve Kendrick from the Information & Statistics Division (Edinburgh).


References


References


References


Appendix I

Databases searched and search strategies

Databases searched

• MEDLINE (OVID Technologies), 1993–1999
• CINAHL (OVID Technologies), 1993 to July 1999
• EMBASE (BIDS, via OVID interface), 1993–1999
• HealthSTAR (OVID Technologies), 1993–1999

Surgical wound infection: prospective studies

The number of publications retrieved by each of the search terms in the following searches is given in brackets at the end of the term.

MEDLINE and HealthSTAR

The number of publications retrieved from MEDLINE is given first, followed by that from HealthSTAR.

1. exp prospective studies [123,516; 103,261]
2. exp cohort studies [372,855; 271,807]
3. exp longitudinal studies [354,688; 254,160]
4. exp follow-up studies [226,796; 149,371]
5. prospective stud$.tw [40,249; 30,492]
6. cohort stud$.tw [10,359; 9621]
7. longitudinal stud$.tw [13,103; 8979]
8. follow-up stud$.tw [4760; 10,253]
9. or/1–8 [394,995; 283,968]
10. exp surgical wound infection/ [16,652; 6353]
11. exp wound infection/ [22,367; 7660]
12. postoperative wound infection.tw [452; 238]
13. surgical wound infection.tw [307; 199]
14. wound infection.tw [4882; 2841]
15. wound infec$.tw [7003; 4031]
16. wound infection adj10 surgery.tw [585; 377]
17. or/16–18 [25,421; 9744]
18. 9 and 17 [3915; 2633]
19. limit 18 to english [3231; 2158]
20. limit 19 to 1993 to 1999 [1044; 1030]
21. limit 20 to Non-Medline [0; 6]

EMBASE

1. exp prospective studies/ [17,462]
2. exp cohort analysis/ [7649]
3. exp cohort studies/ [7649]
4. exp longitudinal studies/ [4438]
5. exp follow-up studies/ [66,997]
6. prospective stud$.tw [30,770]
7. cohort stud$.tw [9631]
8. longitudinal stud$.tw [8525]
9. follow-up stud$.tw [10,332]
10. or/1–9 [150,589]
11. exp surgical infection/ [1120]
12. exp surgical wound infection/ [1120]
13. exp wound infection/ [5110]
14. exp surgical wound/ [538]
15. postoperative wound infection.tw [245]
16. surgical wound infection.tw [144]
17. wound infection.tw [2825]
18. (wound infection adj10 surgery).tw [350]
19. or/11–18 [8114]
20. 10 and 19 [875]
21. limit 20 to english [796]
22. limit 21 to (yr=1993–1999) [638]

CINAHL

1. exp prospective studies/ [14,578]
2. exp follow-up studies/ [14,578]
3. exp longitudinal studies/ [14,578]
4. exp cohort studies/ [14,578]
5. exp concurrent prospective studies/ [21]
7. exp panel studies/ [56]
8. exp pseudolongitudinal studies/ [0]
9. prospective stud$.tw [1554]
10. cohort stud$.tw [901]
11. longitudinal stud$.tw [1077]
12. follow-up stud$.tw [577]
13. panel stud$.tw [47]
14. or/1–13 [15,694]
15. exp wound infection/ [1224]
16. exp surgical wound infection/ [685]
17. postoperative wound infection.tw [15]
18. wound infection.tw [211]
19. wound infec$.tw [323]
20. wound infect$ adj10 surgery.tw [25]
21. or/15–20 [1323]
22. 14 and 21 [119]
23. limit 24 to english [117]
24. limit 25 to (yr=1993–1999) [47]

Cochrane Library

1. SURGICAL-WOUND-INFECTION*.ME [1502]
2. (SURGICAL & WOUND & INFECTION) [2060]
Appendix 1

3. PROSPECTIVE-STUDIES*.ME [25,057]
4. COHORT-STUDIES*.ME [37,303]
5. LONGITUDINAL-STUDIES*.ME [36,581]
6. FOLLOW-UP-STUDIES*.ME [13,781]
7. (#1 OR #2) [2106]
8. (#3 OR #4 OR #5 OR #6) [39,182]
9. ORTHOPEDICS*.ME [193]
10. BURNS*.ME [436]
11. (#9 OR #10) [692]
12. (#7 AND #8) [909]
13. (#12 NOT #11) [894]
14. Limit to [1993–1999] [311]

Surgical wound infection: measurement and validation studies

The number of publications retrieved by each of the search terms in the following searches is given in brackets at the end of the term.

**MEDLINE and HealthSTAR**

The number of publications retrieved from MEDLINE is given first, followed by that from HealthSTAR.

1. exp surgical wound infection/ [16,652; 6353]
2. exp wound infection/ [22,367; 7660]
3. postoperative wound infection.tw [452; 238]
4. surgical wound infection.tw [307; 199]
5. wound infection.tw [4882; 2841]
6. wound infec$.tw [7003; 4031]
7. wound infection adj10 surgery.tw [585; 377]
8. or/1–7 [25,421; 9744]
9. exp reproducibility of results/ [55,155; 50,510]
10. exp “sensitivity and specificity”/ [92,282; 56,854]
11. exp predictive value of tests/ [33,626; 32,117]
12. validity.tw [31,009; 19,062]
13. reliability.tw [30,233; 18,817]
14. sensitivity.tw [19,306; 64,090]
15. specificity.tw [131,651; 36,757]
16. reproducibility.tw [16,719; 8620]
17. test retest.tw [3581; 2788]
18. grading system.tw [1560; 981]
19. scoring system.tw [4101; 2747]
20. or/9–19 [433,406; 185,815]
21. 8 and 20 [528; 328]
22. wound scoring.tw [14; 12]
23. wound grading.tw [1; 1]
24. wound classification.tw [44; 41]
25. exp surgical wound infection/classification [69; 55]
26. exp surgical wound infection/diagnosis [601; 278]
27. exp surgical wound infection/complications [597; 139]
28. or/22–27 [1289; 502]
29. definition.tw [26,699; 13,393]
30. 8 and 29 [79; 59]
31. limit 21 to english [401; 254]
32. limit 31 to 1993–1999 [188; 167]
33. limit 28 to english [881; 365]
34. limit 33 to 1993–1999 [262; 222]
35. limit 30 to english [58; 43]
36. limit 35 to 1993–1999 [23; 23]
37. 32 or 34 or 36 [0; 385]
38. limit 37 to nonmedline [0; 4]

**EMBASE**

1. exp surgical infection/ [1135]
2. exp wound infection/ [5167]
3. exp surgical wound/ [352]
4. postoperative wound infection.tw [245]
5. surgical wound infection.tw [145]
6. wound infection.tw [2856]
7. exp postoperative infection/ [5479]
8. or/1–7 [12,312]
9. exp reproducibility/ [10,217]
10. reproducibility.tw [12,570]
11. exp diagnostic accuracy/ [47,978]
12. sensitivity.tw [129,381]
13. specificity.tw [86,300]
14. validity.tw [22,100]
15. exp *measurement/ [56,055]
16. measurement.tw [88,321]
17. face validity.tw [244]
18. exp reliability/ [13,262]
19. reliability.tw [20,906]
20. exp observer variation/ [2398]
21. intraobserver reliability.tw [75]
22. intrarater reliability.tw [122]
23. exp accuracy/ [16,963]
24. or/9–23 [393,257]
25. 8 and 24 [363]
26. limit 25 to (english language and yr=1993–1999) [216]

**CINAHL**

1. exp wound infection/ [1240]
2. exp surgical wound infection/ [690]
3. postoperative wound infection.tw [15]
4. wound infection.tw [220]
5. wound infec$.tw [332]
6. wound infect$ adj10 surgery.tw [41]
7. or/1–7 [1344]
8. exp reproducibility of results/ [737]
9. exp criterion-related validity/ [1082]
10. exp reliability and validity/ [336]
11. exp validity/ [7619]
12. exp concurrent validity/ [378]
13. external validity/ [235]
14. concurrent validity/ [378]
15. consensual validity/ [382]
16. predictive validity/ [12]
17. discriminant validity/ [174]
18. exp reliability/ [10,280]
19. exp test-retest reliability/ [2028]
20. interrater reliability/ [3334]
21. intrarater reliability/ [775]
22. exp clinical assessment tools/ [9771]
23. or/8–22 [23,356]
24. 7 and 23 [58]
25. wound scoring.tw [2]
26. wound grading.tw [0]
27. wound classification.tw [21]
28. or/25–27 [23]
29. definition.tw [2082]
30. 7 and 29 [4]
31. limit 24 to english [57]
32. limit 28 to english [23]
33. limit 30 to english [3]
34. or/31–33 [82]

Cochrane Library
1. SURGICAL-WOUND-INFECTION*.ME [1502]
2. (SURGICAL and WOUND and INFECTION) [2060]
3. ACCURACY [2372]
4. VALIDITY [4886]
5. REPRODUCIBILITY OF RESULTS*.ME [1994]
6. DIAGNOSTIC-ERRORS*.ME [900]
7. FALSE-NEGATIVE-REACTIONS*.ME [155]
8. FALSE-POSITIVE-REACTIONS*.ME [206]
9. OBSERVER-VARIATION*.ME [544]
10. DIAGNOSIS*.ME [69,023]
11. (#1 OR #2) [2060]
12. (Or #3–10) [74,001]
13. (#10 AND #11) [268]
14. (SURGICAL and (WOUND and (INFECTION and DIAGNOSIS))) [100]
15. (SURGICAL and (WOUND and (INFECTION and CLASSIFICATION))) [41]
16. (WOUND and SCORING) [66]
17. (WOUND and GRADING) [21]
18. (OR #14–17) [202]
19. #18 AND #13 [55]
20. Limit [1993 to 1999] [39]

MEDLINE and HealthSTAR
The number of publications retrieved from MEDLINE is given first, followed by that from HealthSTAR.
1. exp surgical wound infection/ [16,652; 6353]
2. exp wound infection/ [22,367; 7660]
3. postoperative wound infection.tw [452; 238]
4. surgical wound infection.tw [307; 199]
5. wound infection.tw [4882; 2841]
6. wound infec$.tw [7003; 4031]
7. wound infection adj10 surgery.tw [585; 377]
8. or/1–7 [25,421; 9744]
9. exp population surveillance/ [14,546; 12,313]
10. exp sentinel surveillance/ [397; 443]
11. exp data collection/ [432,042; 358,352]
12. exp registries/ [15,174; 11,777]
13. exp medical records/ [255,838; 20,256]
14. exp medical record linkage [859; 843]
15. exp medical records systems, computerised/ [3411; 6029]
16. exp patient discharge/ [6756; 6150]
17. exp outpatients/ [2076; 1754]
18. exp ambulatory surgical procedures/ [4884; 4346]
19. population surveillance.tw [66; 49]
20. wound infection surveillance.tw [22; 16]
21. surveillance.tw [26,010; 17,056]
22. or/9–21 [459,154; 375,953]
23. 8 and 22 [1889; 1678]
24. limit to English language [1551; 1374]
25. limit 24 to year=1993–1999 [943; 940]
26. limit to non-medline [0; 22]

CINAHL
1. exp surgical wound infection/ [690]
2. exp wound infection/ [1240]
3. postoperative wound infection.tw [15]
4. surgical wound infection.tw [37]
5. wound infection.tw [220]
6. wound infec$.tw [332]
7. wound infection adj10 surgery.tw [41]
8. or/1–7 [1344]
9. population surveillance.tw [2]
10. exp disease surveillance/ [1468]
11. sentinel surveillance.tw [7]
12. exp data collection/ [50,099]
13. exp registries, disease/ [363]
14. exp medical records/ [9472]
15. exp medical record linkage [22]
16. exp patient discharge/ [3786]
17. exp outpatients/ [10,806]
18. exp ambulatory surgery/ [11376]
19. after care/ [935]
20. wound infection surveillance.tw [6]
21. surveillance.tw [1759]
22. or/9–21 [68,532]
23. limit 23 to 24 to (yr=1993–yr=1999) [199]

**EMBASE, 1993–1999**
1. exp surgical infection/ [1120]
2. exp wound infection/ [5110]
3. surgical wound infection.tw [144]
4. postoperative wound infection.tw [245]
5. or/1–4 [6162]
6. population surveillance.tw [42]
7. surveillance.tw [17,068]
8. monitoring.tw [75,099]
9. exp hospital discharge/ [3740]
10. hospital discharge.tw [3404]
11. post discharge surveillance.tw [7]
12. postdischarge surveillance.tw [8]
13. exp ambulatory surgery/ [2071]
14. ambulatory surgery.tw [634]
15. wound infection surveillance.tw [12]
16. or/6–15 [98,803]
17. and/5 and 16 [296]
18. limit 16 to English language [259]
19. limit 16 to (yr=1993–yr=1999) [166]

**Cochrane Library**
1. POPULATION-SURVEILLANCE*.ME [141]
2. PATIENT-DISCHARGE*.ME [326]
3. (POSTDISCHARGE and SURVEILLANCE) [3]
4. (PATIENT and DISCHARGE) [1777]
5. AMBULATORY-SURGICAL-PROCEDURES*.ME [692]
6. SURGICAL-WOUND-INFECTION*.ME [1502]
7. (SURGICAL and WOUND and INFECTION) [2060]
8. (#1 OR #2 OR #3 OR #4 OR #5) [2494]
9. (#6 OR #7) [2060]
10. (#8 and #9) [64]
11. Limit to [1993 to 1999] [43]

**Anastomotic leak**

The number of publications retrieved by each of the search terms in the following searches is given in brackets at the end of the term.

**MEDLINE and HealthSTAR**

The number of publications retrieved from MEDLINE is given first, followed by that from HealthSTAR.

1. exp anastomosis, surgical/ [3240; 13,782]
2. exp anastomosis, surgical/ae,cl,mo [5817; 2726]
3. anastomotic leak.tw [511; 361]
4. anastomotic leakage.tw [705; 413]
5. anastomotic breakdown.tw [55; 25]
6. anastomotic dehiscence.tw [274; 153]
7. anastomo$.tw [32,613; 11,409]
8. leak.tw [7226; 9226]
9. leak$.tw [27,439; 9820]
10. breakdown.tw [18,502; 2866]
11. dehiscence.tw [2726; 1470]
12. exp cholecystostomy/ [334; 263]
13. exp choledochostomy/ [589; 396]
14. exp gastroenterostomy/ [2239; 543]
15. exp jejunoileal bypass/ [394; 141]
16. exp pancreaticojejunosotomy/ [291; 226]
17. exp portoenterostomy, hepatic/ [188; 148]
18. exp prospective studies [123,516; 103,261]
19. exp cohort studies [372,855; 271,807]
20. exp longitudinal studies [354,688; 254,160]
21. exp follow-up studies [226,796; 149,371]
22. prospective stud$.tw [40,249; 30,492]
23. cohort stud$.tw [10,359; 9621]
24. longitudinal stud$.tw [13,103; 8979]
25. follow-up stud$.tw [18,184; 10,255]
26. or/1–2 [32,460; 13,782]
27. or/3–10 [77,346; 23,434]
28. 26 and 27 [9657; 4902]
29. or/12–17 [394,995; 283,968]
30. 28 or 30 [9657; 4902]
31. and/5 and 16 [296]
32. limit 16 to English language [259]
33. limit 16 to (yr=1993–yr=1999) [166]

**EMBASE**
1. exp anastomosis leakage/ [825]
2. exp *anastomosis dehiscence/ [71]
3. exp *anastomosis leakage/ [208]
4. anastomotic leak$.tw [1057]
5. anastomotic leakage.tw [480]
6. or/1–5 [1533]
7. limit 6 to english [1330]
8. limit 7 to (yr=1993–1999) [855]

**CINAHL**
1. exp anastomosis surgical/ [190]
2. anastomotic leak.tw [5]
3. anastomotic leakage.tw [2]
4. anastomotic breakdown.tw [0]
5. anastomotic dehiscence.tw [0]
6. anastomol$.tw [76]
7. leak.tw [128]
8. leak$.tw [392]
9. breakdown.tw [423]
10. dehiscence.tw [53]
11. or/1–6 [193]
Deep vein thrombosis

The number of publications retrieved by each of the search terms in the following searches is given in brackets at the end of the term.

MEDLINE and HealthSTAR

The number of publications retrieved from MEDLINE is given first, followed by that from HealthSTAR.

1. exp venous thrombosis/ [21,890; 6589]
2. exp venous thrombosis/di [4236; 1339]
3. exp venous thrombosis/us [758; 661]
4. exp thrombophlebitis [17,550; 4426]
5. exp thrombophlebitis/di [3349; 895]
6. venous thrombosis/ [1749; 1307]
7. venous thrombosis.tw [8367; 3842]
8. deep vein thrombosis.tw [3936; 2191]
9. dvt.tw [1919; 1376]
10. thromboembolism.tw [7116; 3741]
11. or/1–10 [31,526; 11,493]
12. exp sensitivity and specificity/ [64,446; 56,854]
13. exp reproducibility of results/ [55,115; 50,510]
14. exp predictive value of tests/ [33,626; 32,117]
15. test retest.tw [3581; 2788]
16. validity.tw [31,009; 19,062]
17. reliability.tw [30,253; 18,817]
18. sensitivity.tw [193,067; 64,090]
19. specificity.tw [131,651; 36,757]
20. reproducibility.tw [16,719; 8620]
21. or/12–20 [427,664; 183,098]
22. 11 and 21 [1373; 1048]
23. limit 22 to English (1993–1999) [1109; 573]
24. limit to nonmedline [0; 7]

CINAHL

1. exp venous thrombosis/ [666]
2. exp venous thrombosis/di [141]
3. exp venous thrombosis/us [43]
4. exp thrombophlebitis [666]
5. exp thrombophlebitis/di [141]
6. venous thrombosis/ [208]
7. venous thrombosis.tw [260]
8. deep vein thrombosis.tw [232]
9. dvt.tw [154]
10. thromboembolism.tw [226]
11. or/1–10 [961]
12. exp sensitivity and specificity/ [1892]
13. exp reproducibility of results/ [737]
14. exp predictive value of tests/ [743]
15. test retest.tw [767]
16. validity.tw [3981]
17. reliability.tw [4143]
18. sensitivity.tw [2694]
19. specificity.tw [1096]
20. reproducibility.tw [268]
21. or/12–20 [10,955]
22. 11 and 21 [31]
23. limit 22 to English [31]

EMBASE

1. exp deep vein thrombosis/ [6985]
2. exp deep vein thrombosis/di [211]
3. exp *deep vein thrombosis/di [1729]
4. deep vein thrombosis.tw [2993]
5. exp *thromboembolism/di [9529]
6. exp *thrombophlebitis/di [203]
7. exp reproducibility/ [10,217]
8. reproducibility.tw [12,570]
9. exp diagnostic accuracy/ [47,978]
10. sensitivity.tw [129,381]
11. specificity.tw [86,300]
12. or/1–6 [15,076]
13. or/7–11 [233,121]
14. 12 and 13 [1450]
15. limit 14 to (english language/ yr=1993–1999) [873]

Cochrane Library

1. (DEEP and (VEIN and THROMBOSIS)) [1218]
2. THROMBOEMBOLISM [1093]
3. THROMBOEMBOLISM*.ME [528]
4. DVT [544]
5. (VENOUS and THROMBOSIS) [1333]
6. THROMBOPHLEBITIS [1258]
7. THROMBOPHLEBITIS*.ME [966]
8. OR #1 to #7 [3312]
9. DIAGNOSIS*.ME [69,023]
10. MEASUREMENT [9643]
11. ACCURACY [2372]
12. SENSITIVITY [10,889]
13. SPECIFICITY [4082]
14. SENSITIVITY-AND-SPECIFICITY*.ME [3408]
15. DIAGNOSTIC-TECHNIQUES-AND-PROCEDURES*.ME [48,412]
16. DIAGNOSTIC and IMAGING [1565]
17. (#8 AND #17) [971]
18. Limit to [1993 to 1999] [522]

**Mortality-related monitoring systems**

The number of publications retrieved by each of the search terms in the following searches is given in brackets at the end of the term.

**MEDLINE and HealthSTAR**
The number of publications retrieved from MEDLINE is given first, followed by that from HealthSTAR.

1. exp surgery/ [15,346; 6968]
2. exp colorectal surgery/ [549; 315]
3. exp thoracic surgery/ [3831; 1503]
4. exp surgical procedures, operative/ [1,085,562; 380,452]
5. surgery.tw [288,049; 143,160]
6. surgical procedur$.tw. [24,634; 13,329]
7. or/1–6 [1,240,093; 444,969]
8. exp mortality/ [112,599; 84,217]
9. exp hospital mortality/ [3832; 4109]
10. exp surgical procedures, operative/mo [14,749; 9770]
11. exp surgical procedures, laparoscopic/mo [131; 109]
12. exp surgical procedures, endoscopic/mo [237; 187]
13. mortality.tw. [139,490; 82,513]
14. operative death.tw. [820; 501]
15. surgical mortality.tw. [914; 549]
16. postoperative death.tw. [677; 416]
17. postoperative mortality.tw. [1916; 1259]
18. or/8–17 [227,637; 145,651]
19. exp population surveillance/ [14,546; 12,313]
20. exp records/ [52,321; 38,189]
21. exp registries/ [15,174; 11,777]
22. exp information systems/ [35,102; 37,502]
23. exp medical audit/ [5628; 4738]
24. monitoring system.tw. [1778; 1059]
25. information system.tw. [3103; 2924]
26. mortality surveillance.tw. [108; 93]
27. cepod.tw. [8; 8]
28. ncepod.tw. [20; 22]
29. surgical audit.tw. [135; 111]
30. surgery audit.tw. [23; 20]
31. critical incident.tw. [266; 249]
32. confidential inquiry.tw. [30; 19]
33. confidential enquiry.tw. [87; 81]
34. or/19–33 [101,362; 884,378]
35. 7 and 18 [73,784; 52,220]
36. 7 and 34 [9546; 8224]
37. 18 and 34 [9788; 8686]
38. 35 and 34 [2189; 2037]
39. limit 38 to english language [1995; 1862]
40. limit 39 to yr=1993–1999 [1285; 1304]
41. limit to non-medline [0; 28]

**EMBASE**

1. exp *surgical mortality/ [302]
2. surgical mortality.tw [471]
3. exp *mortality/ [11,489]
4. mortality.tw [90,392]
5. hospital mortality.tw [3441]
6. operative mortality.tw [3258]
7. surgical death.tw [45]
8. or/1–7 [93,426]
9. exp *register/ [309]
10. registry.tw [8418]
11. registries.tw [1794]
12. exp *information system/ [3005]
13. monitoring system.tw [374]
14. information system.tw [1997]
15. population surveillance.tw [42]
16. surveillance.tw [17,264]
17. ncepod.tw [18]
18. cepod.tw [7]
19. confidential enquiry.tw [59]
20. confidential inquiry.tw [19]
21. or/9–20 [31,146]
22. 8 and 21 [3382]
23. limit 22 to english language [3103]
24. limit 23 to yr=1993–1999 [1985]

**CINAHL, 1993–1999**

1. exp hospital mortality/ [160]
2. exp mortality/ [3693]
3. mortality.tw [5550]
4. surgical mortality.tw [5]
5. hospital mortality.tw [105]
6. postoperative death.tw [3]
7. surgical death.tw [0]
8. or/1–7 [7858]
9. exp disease registries/ [363]
10. exp information systems/ [9955]
11. exp disease surveillance/ [1468]
12. information system.tw [560]
13. monitoring system.tw [149]
14. mortality surveillance.tw [10]
15. cepod.tw [0]
16. ncepod.tw [0]
17. confidential enquiry.tw [37]
18. confidential inquiry.tw [3]
19. surgical audit.tw [2]
20. or/9–19 [11,978]
21. 8 and 20 [264]
22. limit 21 to yr=1993–1999 [207]
23. limit 22 to english [204]

**Cochrane Library**
1. (SURGICAL and MORTALITY) [2026]
2. (SURGICAL and DEATH) [745]
3. (POSTOPERATIVE and DEATH) [578]
4. (#1 OR #2 OR #3) [2638]
5. (INFORMATION and SYSTEM) [2201]
6. INFORMATION SYSTEMS*.ME [443]
7. REGISTRIES [222]
8. SURVEILLANCE [1020]
9. (SURVEILLANCE and SYSTEM) [129]
10. (MONITORING and SYSTEM) [914]
11. (#5 OR #6 OR #7 OR #8 OR #9 OR #10) [4363]
12. (#4 AND #11) [253]
Appendix 2

Statistical assessment criteria and data extraction forms

Assessment of diagnostic utility and accuracy

The comparison of a definition and/or measurement of a surgical adverse event was quantified using the following assessments. A disease may be present or absent and a test result either positive or negative, and therefore, when expressed in a contingency table, four combinations of disease status and test result are possible. The statistics that can be calculated from the contingency table (Table 20) are sensitivity, specificity, positive predictive value and negative predictive value.

- The **sensitivity** is measured by the number of true-positive results correctly identified by the definition or measurement technique out of all the true-positive results identified by the gold standard \((a/a + c)\).
- The **specificity** is measured by the number of true-negative results correctly identified by the definition or measurement technique out of all the true-negative results identified by the gold standard \((d/b + d)\).
- The **positive predictive value** (PPV) is measured by the number of true-positive results correctly identified by the definition or measurement technique out of all the positive results identified by the definition or measurement method being tested or compared \((a/a + b)\).
- The **negative predictive value** (NPV) is measured by the number of true-negative results correctly identified by the definition or measurement technique out of all the negative results identified by the definition or measurement method being tested or compared \((d/c + d)\).

Assessment of inter- and intra-reliability: statistical measures of agreement

**Categorical variables, \(\kappa\)**

The value of \(\kappa\) can be used to assess the extent of agreement between two raters, or between two alternative classification or diagnostic methods. The parameter \(\kappa\) measures the chance-corrected proportional agreement.\(^{395}\) It can take values up to 1 (perfect agreement), with a value of 0 indicating no agreement better than chance and a negative value indicating that agreement is worse than chance. The interpretation of \(\kappa\) follows the guidelines given in Table 21.\(^{396}\)

**Quantitative data**

**Repeatability (intra-observer reliability)**

Repeatability refers to the variability of repeated measurements taken under similar conditions (e.g. the same observer). We would expect that 95% of the differences in measurements would be less than two standard deviations. The **coefficient of repeatability** is therefore defined as two times the standard deviation of the differences.\(^{13}\)

<table>
<thead>
<tr>
<th>Table 21</th>
<th>Interpretation of (\kappa)*</th>
</tr>
</thead>
<tbody>
<tr>
<td>(\kappa)</td>
<td>Strength of agreement</td>
</tr>
<tr>
<td>&lt; 0.2</td>
<td>Poor</td>
</tr>
<tr>
<td>0.21–0.4</td>
<td>Fair</td>
</tr>
<tr>
<td>0.41–0.6</td>
<td>Moderate</td>
</tr>
<tr>
<td>0.61–0.8</td>
<td>Good</td>
</tr>
<tr>
<td>0.81–1.0</td>
<td>Very good</td>
</tr>
</tbody>
</table>

* Adapted from Landis and Koch\(^{443}\)

**TABLE 20** Contingency table of disease status and test results

<table>
<thead>
<tr>
<th>Disease present</th>
<th>Disease absent</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>Test positive</td>
<td>True-positive results, (a)</td>
<td>False-positive results, (b)</td>
</tr>
<tr>
<td>Test negative</td>
<td>False-negative results, (c)</td>
<td>True-negative results, (d)</td>
</tr>
<tr>
<td>Total</td>
<td>Total true-positive results ((a + c))</td>
<td>Total true-negative results ((b + d))</td>
</tr>
</tbody>
</table>
**Reproducibility (inter-observer reliability)**

Reproducibility refers to the variability of repeated measurements taken under different conditions (e.g., different observers, different methods of measurement). As with the assessment of repeatability, an indicator of the reproducibility is assessed using the standard deviation to produce limits of agreement. Limits of agreement are constructed from the mean difference plus two standard deviations.\(^\text{13}\)

**Data extraction forms**

The data extraction form used is shown in *Table 22.*
TABLE 22  Data extraction form

<table>
<thead>
<tr>
<th>Field</th>
<th>Data to be completed</th>
</tr>
</thead>
<tbody>
<tr>
<td>SECTION 1  (this section printed on every form)</td>
<td></td>
</tr>
<tr>
<td>ID number</td>
<td>Reference Manager identification number</td>
</tr>
<tr>
<td>Authors</td>
<td>All authors</td>
</tr>
<tr>
<td>Journal details</td>
<td>Journal title, volume, publication year, chapter, page numbers</td>
</tr>
<tr>
<td>Title</td>
<td>Article title</td>
</tr>
<tr>
<td>Aim</td>
<td>Aim of the study as specified in the article text rather than the abstract</td>
</tr>
<tr>
<td>Study design</td>
<td>Prospective, measurement/validation study, surveillance/information system, other (specify)</td>
</tr>
<tr>
<td>Surgical specialty</td>
<td>General; ear, nose and throat; orthopaedic; gynaecological; cardiac; vascular; paediatric; eye; neurological; head and neck; urological; other (specify)</td>
</tr>
<tr>
<td>Procedure</td>
<td>Specify procedure or multiple procedures</td>
</tr>
<tr>
<td>Surgical adverse event</td>
<td>Surgical wound infection, anastomotic leak, deep vein thrombosis, mortality, other (specify)</td>
</tr>
<tr>
<td>Search details</td>
<td>Identifier from MEDLINE, EMBASE, CINAHL, HealthSTAR, Cochrane Library, reference checking, other (specify)</td>
</tr>
<tr>
<td>SECTION 2  Prospective studies  (surgical wound infection, anastomotic leak)</td>
<td></td>
</tr>
</tbody>
</table>
| Definition of adverse event | Has a definition of the adverse event been given? (yes/no)  
Give the definition and list any references to it |
| Surgical wound infection | Has a scoring system been used to grade severity of infection? (yes/no)  
Describe the system and list references to it  
Has a wound classification been given?  
Has a risk index been used?  
Further comments |
| Anastomotic leak | How has anastomotic leak been diagnosed?  
(a) Clinical, with or without radiological confirmation. Describe  
(b) Routine imaging (regardless of clinical signs). Describe  
(c) Prolonged leakage from drain. Describe |
| SECTION 3  Validation studies | |
| Definition of adverse event | Has a definition of the adverse event been given? (yes/no)  
Give definition and list any references to it |
| Validity | How has the adverse event been measured? (subjective assessment/objective assessment/both) |
| Face validity | On the face of it, does the measure/grading system/test appear to be relevant, reasonable, unambiguous and clear? |
| Objective assessment | List the test undertaken to assess the adverse event (e.g. swab, culture, radiological investigation) |
| Accuracy | Has the adverse event been compared with a standard? Describe the standard  
Is there a report of agreement with the standard? |
| Sensitivity | Has sensitivity been reported? (ability of gradations in score to reflect change, probability of correctly identifying the adverse event) |

continued
TABLE 22 contd Data extraction form

<table>
<thead>
<tr>
<th>Field</th>
<th>Data to be completed</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>SECTION 3 Validation studies contd</strong></td>
<td></td>
</tr>
</tbody>
</table>
| Specificity | Has specificity been reported? (measure of probability of correctly identifying a non-affected case with the measure)  
Has the level of agreement been given? |
| Reliability | The extent to which the measure is consistent and minimises random error (repeatability) and inter-/intra-observer variation  
Did more than one person assess the event? Describe  
Was there agreement between assessors? Describe  
If so, how were discrepancies resolved? Describe  
Did the same person assess the event at two or more points in time? |
| Practicality | Do the authors refer to the practicality or feasibility of the measure?  
Is the instrument easy to administer and process?  
Consider time and resources |
| **SECTION 4 Surveillance/monitoring studies** | |
| Type of surveillance | Continuous or intermittent, all surgical procedures or selective? |
| Coverage | Participating hospitals and departments |
| Extent of coverage | Denominator  
Are denominator data collected or available? |
| Data collection | Process by which data are extracted, by whom and how  
Give details of staff, process, methods and data collection forms (when applicable) |
| Risk analysis | Is risk stratification used?  
Give index/indices used |
| Validation | List details of validation (e.g. accuracy of case ascertainment, accuracy of data input) |
| Feedback | Describe feedback to users, format of output and dissemination |
| Other | Describe other features of system not covered by the headings above (e.g. cost of system, any description of impact) |
### Appendix 3
Summary of included studies

<table>
<thead>
<tr>
<th>Study</th>
<th>Study design, sample size, aim</th>
<th>Definition, grading system, cited references</th>
<th>Data collection, postdischarge surveillance</th>
</tr>
</thead>
<tbody>
<tr>
<td>Abramov et al., 1996, Israel</td>
<td>RCT of antibiotic prophylaxis (n = 99) in umbilical and incisional hernia repair</td>
<td><strong>Definition:</strong> wound infection defined as erythema and drainage from a wound or a wound with a purulent discharge alone. A probable wound infection was considered to be present if erythema extended at least 2 cm from the wound in any direction</td>
<td><strong>Hospital:</strong> sought daily for duration of admission <strong>Community:</strong> all patients examined 1 week postoperatively and invited for re-examination 3 weeks later</td>
</tr>
<tr>
<td></td>
<td><strong>Aim:</strong> to ascertain whether preoperative antibiotic prophylaxis would reduce wound infection rates after umbilical and incisional hernia repair</td>
<td><strong>Reference to definition:</strong> Sheretz.33 Consensus paper on the surveillance of SWIs; Glenister et al.; Platt and co-workers124</td>
<td></td>
</tr>
<tr>
<td></td>
<td><strong>Grading system:</strong> none</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Anon., 1996, Multicentre prospective study of mediastinitis of cardiac surgery (n = 1830)</td>
<td></td>
<td><strong>Hospital:</strong> sternal wound assessed on daily basis during patient’s stay in the cardiac surgical unit. If patient discharged before postoperative day 8, surveillance continued in new location until at least that date <strong>Community:</strong> patients not systematically surveyed beyond day 8</td>
<td></td>
</tr>
<tr>
<td></td>
<td><strong>Aim:</strong> to ascertain the incidence of deep sternal wound infection and identify high-risk patients or procedures</td>
<td><strong>Definition:</strong> 1988 CDC31 Further definition of DSWI given as: an infection involving tissues of the spaces beneath the subcutaneous tissue. In addition, DSWI must fit at least one of the following criteria: (a) an organism is isolated from culture of mediastinal tissue or fluid; (b) evidence of mediastinitis is seen during operation or on histopathological examination; or (c) fever (&gt; 38°C), chest pain, or sternal instability is present and there is either purulent drainage from the mediastinum or an organism isolated from blood culture or from culture of drainage of the mediastinal area. DSWI was categorised as osteomyelitis when infection was limited to the sternal bone and as mediastinitis when infection included the mediastinal area</td>
<td></td>
</tr>
<tr>
<td></td>
<td><strong>Grading system:</strong> none</td>
<td><strong>Reference to definition:</strong> Garner and co-workers31</td>
<td></td>
</tr>
<tr>
<td>Barber et al., Prospective study of cancer surgery (n = 1283)</td>
<td></td>
<td><strong>Hospital:</strong> observer’s evaluations of SSI validated against observations of an attending surgeon. Examined on postoperative day 3 and every 3 days thereafter. Observation schedule increased to every 48 hours, or more frequently at discretion of observer, or when surgical sites were observed to be infected <strong>Community:</strong> early discharge patients were interviewed by telephone, mailed a brief questionnaire or observed in the outpatient clinic</td>
<td></td>
</tr>
<tr>
<td></td>
<td><strong>Aim:</strong> to identify the rate of SSI and risk factors for SSI in patients with cancer and to evaluate antibiotic use patterns on surgical oncology services</td>
<td><strong>Definition:</strong> 1992 CDC12 Any surgical site that was graded 3+ or 4+ was considered infected. Pus exuding from puncture sites of sutures or staples only (i.e. suture abscesses) without other over signs (i.e. cellulitis &gt; 1 cm) was not considered evidence of SSI. Wounds culture positive in the presence of physical signs of infection (i.e. signs of inflammation, fever) were also graded as infected, as were surgical sites considered to be infected based on this study’s criteria</td>
<td></td>
</tr>
<tr>
<td></td>
<td><strong>Grading system:</strong> Coit142</td>
<td><strong>Reference to definition:</strong> Horan and co-workers32</td>
<td></td>
</tr>
</tbody>
</table>

**DSWI, deep sternal wound infection; RCT, randomised controlled trial; SSI, surgical site infection; SWI, surgical wound infection**

*continued*
## TABLE 23 contd Prospective studies of surgical wound infection

<table>
<thead>
<tr>
<th>Study</th>
<th>Study design, sample size, aim</th>
<th>Definition, grading system, cited references</th>
<th>Data collection, postdischarge surveillance</th>
</tr>
</thead>
<tbody>
<tr>
<td>Bellchambers et al., 1999, UK</td>
<td>RCT of two draping systems in cardiac surgery (n = 593)</td>
<td>Definition: sternal and leg wounds scored by ASEPSIS. A clinically significant wound infection had a score of 21–30; a deep wound infection, usually associated with sternal click, scored 31–40; bone infection scored &gt; 41</td>
<td>Hospital: blinded assessment assessed daily for 5 of 7 days postoperatively Community: information at 6 weeks and 3 months postoperatively, collected by telephone or at clinic according to protocol</td>
</tr>
<tr>
<td><strong>Aim</strong>: a prospective audit of wound infections in a tertiary referral centre for cardiac surgery and to evaluate the clinical use of an established wound scoring system (ASEPSIS)</td>
<td><strong>Reference to definition</strong>: Wilson et al.140</td>
<td><strong>Grading system</strong>: ASEPSIS140</td>
<td></td>
</tr>
<tr>
<td>Bencini et al., 1994, Italy</td>
<td>RCT of topical gel in cutaneous centrofacial lesions (n = 673)</td>
<td>Definition: infection suspected on basis of both: (1) the presence of at least two of the following clinical criteria — inflammation with serous discharge or purulence or presence of fibrinous debris, tenderness and warmth; and (2) the presence of leucocytes and bacteria on microscopic examination of exudate with methylene blue stain. Infection confirmed by culture in every case</td>
<td>Hospital: independent evaluation by member of surgical staff and blinded evaluator. Wounds dressed and assessed by same physician on days 2, 4, and 6 postoperatively Community: not assessed. Final evaluation on day 6 postoperatively</td>
</tr>
<tr>
<td><strong>Aim</strong>: to analyse the effectiveness of topical benzoyl peroxide gel for the prevention of surgical skin wound infections in the centrofacial area</td>
<td><strong>Reference to definition</strong>: none</td>
<td><strong>Grading system</strong>: none</td>
<td></td>
</tr>
<tr>
<td>Bold et al., 1998, USA</td>
<td>RCT of antibiotic prophylaxis in axillary lymph node dissection ± mastectomy (n = 200)</td>
<td>Definition: erythema and induration of the surgical wound that required the initiation of antibiotics, purulent drainage from the incision, or systemic symptoms of an infection (i.e. fever with malaise or anorexia) in the absence of any other site of infection</td>
<td>Hospital: many patients discharged within 23 hours Community: all patients followed up in the outpatient clinic and monitored for infection. A research nurse contacted patients and the referring physician for wound follow-up at 4 weeks postoperatively</td>
</tr>
<tr>
<td><strong>Aim</strong>: to determine whether a single dose of preoperative cefalosporin could decrease the incidence of postoperative wound complications following axillary lymph node dissection</td>
<td><strong>Reference to definition</strong>: none</td>
<td><strong>Grading system</strong>: none</td>
<td></td>
</tr>
<tr>
<td>Brown et al., 1996, USA</td>
<td>Prospective study in cardiac surgery (n = 1717)</td>
<td>Definition: 1992 CDC12 Reference to definition: Horan and co-workers,12 Sherertz13</td>
<td>Hospital: wound assessed daily by infection control staff (nurse epidemiologist) independent from surgical staff. Data collected as per NNIS system and an Infection Surveillance Record started when an infection was suspected Community: arrangements made for follow-up reports from doctor’s office postdischarge</td>
</tr>
<tr>
<td><strong>Aim</strong>: to determine the operative wound infection rates occurring in a series of 1717 cardiac operations performed under direct ultraviolet C radiation</td>
<td><strong>Grading system</strong>: none</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Byrne et al., 1994, Scotland</td>
<td>RCT of whole body disinfection (n = 3733) in clean and clean–contaminated procedures</td>
<td>Definition: infection defined as an ASEPSIS score &gt; 10 and/or a discharge of pus. For the remainder of the study (after discharge) wound infection was defined as a discharge of pus from the wound either spontaneously or after incision</td>
<td>Hospital: inspection and scoring of wounds daily until discharge Community: at the outpatient clinic at 6 weeks information about wound was recorded on a standard form. If primary healing had not occurred, details were obtained and the GP/nurse contacted for further information. Patients who failed to attend the clinic were followed up by postal questionnaire and further details obtained from the GP/nurse. Follow-up was 99% complete (70% at clinic, 30% by questionnaire). In total of 922 (27%) patients were fully scored by ASEPSIS, which required that were inpatients for at least 7 days postoperatively</td>
</tr>
<tr>
<td><strong>Aim</strong>: to study the importance of the definition and postdischarge wound surveillance on reported wound infection rates using data from an RCT</td>
<td><strong>Reference to definition</strong>: Ljungqvist121</td>
<td><strong>Grading system</strong>: ASEPSIS90</td>
<td></td>
</tr>
<tr>
<td>RCT, randomised controlled trial</td>
<td></td>
<td></td>
<td>continued</td>
</tr>
<tr>
<td>Study</td>
<td>Study design, sample size, aim</td>
<td>Definition, grading system, cited references</td>
<td>Data collection, postdischarge surveillance</td>
</tr>
<tr>
<td>-------</td>
<td>-------------------------------</td>
<td>-----------------------------------------------</td>
<td>---------------------------------------------</td>
</tr>
</tbody>
</table>
| den Hoed et al., 1998, 85 Netherlands | Prospective audit of open and laparoscopic cholecystectomy (n = 637)  
Aim: to determine the incidence of postoperative infections, especially wound infections, after open and laparoscopic biliary surgery, and to assess the bacteriological data on these patients | Definition: used the Mitchell scoring system for the definition  
Reference to definition: used the Mitchell scoring system for the definition  
Grading system: Mitchell et al.122 | Hospital: daily scores for serous exudate, erythema, purulent discharge and separation of deep tissue  
Community: items scored again at follow-up at outpatients. Open cholecystectomy assessed at 3 weeks postoperatively; laparoscopic cholecystectomy assessed at 1 and 3 weeks. Results presented as in-hospital or delayed wound infection if the infection appeared after discharge |
| Fanning et al., 1995, 98 Canada | Prospective randomised study of clean and clean–contaminated surgical procedures (n = 1200)  
Aim: to evaluate three methods for conducting postdischarge SSI surveillance | Definition: purulent discharge from a surgical site with or without a positive culture. Stitch abscesses were excluded if the inflammation and exudate were confined to the suture penetration site  
Reference to definition: none  
Grading system: none | Hospital: no details given as the aim was to determine postdischarge rates  
Community: three methods were used: (1) a questionnaire was given at discharge to be completed and returned at 30 days; (2) surgeons were mailed a follow-up card within 2 weeks of surgery; to be returned after the 4 to 6 week postoperative clinic; and (3) telephone contact 30 days after surgery by the ICP, who administered a standard questionnaire over the telephone. The highest rate of return was from the telephone |
| Fenton-Lee et al., 1994, 86 UK | Prospective audit of multiple surgical procedures (n = 463)  
Aim: to assess the patient acceptability and outcome of day-surgery before and after changes to the service | Definition: as per the wound grading system below  
Reference to definition: as per wound grading system  
Grading system: Cruse436 wound grading system (although this article does not contain a grading system) | Hospital: day-surgery unit  
Community: patients followed up at 1, 7 and 30 days postoperatively by the liaison sister, who documented wound complications using the standard grading system |
| Ferraz et al., 1995, 99 Brazil | Prospective study of general surgical and caesarean procedures (n = 6604)  
Aim: to evaluate the method of postdischarge surveillance for SWI after general surgical and caesarean procedures from 1988 to 1992 | Definition: the presence of pus, with or without confirmation by culture  
Reference to definition: none  
Grading system: none | Hospital: infection commission nurse reviewed all patients daily and observed wounds  
Community: at discharge, patients were instructed to report at a central outpatient clinic on day 8 post-operatively. The postdischarge return rate for general surgery ranged from 68.4% to 84.6% over the 5-year period |
| Gipponi et al., 1993, 100 Italy | RCT of immunoprophylaxis in gastrointestinal surgery (n = 369)  
Aim: to evaluate the effectiveness and tolerability of perioperative prophylaxis with intravenous immunoglobulin in ‘septic-risk’ patients undergoing surgical treatment for gastro-intestinal cancer | Definition: primary wound infections (when the discharge is pus); secondary wound infections (when the first discharge is not pus, but the discharging wound becomes colonised by bacteria from an endogenous sources, i.e. anastomotic dehiscence)  
Reference to definition: Pollock125  
Grading system: none | Hospital: all patients were observed until they were released from hospital or died  
Community: no details given |

RCT, randomised controlled trial; SSI, surgical site infection; SWI, surgical wound infection

continued
<table>
<thead>
<tr>
<th>Study</th>
<th>Study design, sample size, aim</th>
<th>Definition, grading system, cited references</th>
<th>Data collection, postdischarge surveillance</th>
</tr>
</thead>
<tbody>
<tr>
<td>Grant et al., 1995,101 USA</td>
<td>Prospective study of elective colorectal surgery (n = 89)</td>
<td><strong>Definition:</strong> as per the grading system below</td>
<td><strong>Hospital:</strong> wounds assessed daily during inpatient stay <strong>Community:</strong> outpatient examinations performed for 4 weeks or more for close follow-up of wound healing</td>
</tr>
<tr>
<td>Groot and Chappell, 1994,102 Canada</td>
<td>RCT of a scalpel and cautery technique in abdominal and thoracic surgery (n = 672)</td>
<td><strong>Definition:</strong> infected if purulent discharge or if pathogenic organisms were cultured from a wound showing erythema or seroma, or both</td>
<td><strong>Hospital:</strong> assessment done by independent blinded assessor. One of two nurses assessed each wound on Monday, Wednesday and Friday <strong>Community:</strong> at the first clinic visit, at 6 weeks, the surgeon indicated wounds would result in increased wound infection rates</td>
</tr>
<tr>
<td>Hakansson et al., 1993,103 Denmark</td>
<td>RCT of antibiotic prophylaxis in elective colorectal surgery (n = 660)</td>
<td><strong>Definition:</strong> as the discharge of pus from the wound</td>
<td><strong>Hospital:</strong> no details given <strong>Community:</strong> no details given</td>
</tr>
<tr>
<td>Hansen et al., 1996,104 Australia</td>
<td>RCT of open versus laparoscopic appendicectomy (n = 158)</td>
<td><strong>Definition:</strong> purulent discharge from an incision site</td>
<td><strong>Hospital:</strong> no details given <strong>Community:</strong> review at outpatient clinic at 1 and 4 weeks postoperatively. Those who did not return were interviewed by telephone</td>
</tr>
<tr>
<td>Holm et al., 1998,105 Denmark</td>
<td>RCT of wound dressings in abdominal surgery with incisions &gt; 5 cm (n = 73)</td>
<td><strong>Definition:</strong> infection diagnosed by pus, pyrexia and local tenderness</td>
<td><strong>Hospital:</strong> daily inspection until discharge <strong>Community:</strong> cosmetic outcome was only assessed at the final follow-up at 3 months</td>
</tr>
</tbody>
</table>

*RCT, randomised controlled trial*
TABLE 23 contd Prospective studies of surgical wound infection

<table>
<thead>
<tr>
<th>Study</th>
<th>Study design, sample size, aim</th>
<th>Definition, grading system, cited references</th>
<th>Data collection, postdischarge surveillance</th>
</tr>
</thead>
<tbody>
<tr>
<td>Holmes and Readman, 1994, UK</td>
<td>Prospective study of elective inguinal hernia repair as day-surgery (n = 106)</td>
<td>Definition: purulent discharge in or exuding from a wound, or seen on direct examination of the operative site within 30 days after surgical operation. The infection must appear related to the surgical procedure</td>
<td>Hospital: day surgery unit Community: a letter and questionnaire were given to patients to be completed by their GP or nurse (district or practice), who was requested to take wound swab if a SWI was suspected Patients were contacted by telephone at 1 month postoperatively. Researchers contacted the GPs and patients’ notes were inspected by an infection control nurse at the outpatient clinic</td>
</tr>
<tr>
<td>Hopf et al., 1997, USA</td>
<td>Prospective study of general surgical patients (n = 130)</td>
<td>Definition: drainage of purulent material with or without positive culture results</td>
<td>Hospital: operative wounds were inspected daily by the research team during hospitalisation Community: wounds were inspected by the surgeon at the outpatient clinic until day 30 postoperatively</td>
</tr>
<tr>
<td>Israelsson et al., 1996, Sweden</td>
<td>Prospective study of midline laparotomy (n = 467)</td>
<td>Definition: a purulent discharge from the wound (bacterial cultures were not demanded)</td>
<td>Hospital: no details given Community: wound complications were recorded for 12 months; wounds were assessed for incisional hernia. Examination was by the same surgeon apart from for those patients who had left the area, who were examined by a local physician</td>
</tr>
<tr>
<td>Israelsson et al., 1997, Sweden</td>
<td>Prospective observational study of midline laparotomy (n = 1023)</td>
<td>Definition: a purulent discharge from the wound, with or without microbiological culture</td>
<td>Hospital: no details given Community: patients were examined at 12 months by the same surgeon, apart from for those patients who had left the area, who were examined by a local physician</td>
</tr>
<tr>
<td>Jewesson et al., 1996, Canada</td>
<td>RCT of antibiotic prophylaxis in elective biliary tract surgery (n = 167)</td>
<td>Definition: used ASEPSIS and a classification scheme for the definition</td>
<td>Hospital: wounds were assessed daily by hospital and research staff (blinded), until discharge Community: patients were contacted by telephone 30 days after surgery to assess the status of the wounds. If infection was identified or suspected, the surgeon was contacted to determine the nature of the infection and the treatment required</td>
</tr>
</tbody>
</table>

**References:**
- Glenister et al., 28
- Simmons, 433
- Israelsson et al., 105
- Israelsson et al., 106
- Wilson et al., 90

**Grading system:** none except for those patients who had midline laparotomy wounds left the area, who were examined by a local physician.

**Data collection, postdischarge surveillance:**
- Hospital: day surgery unit
- Community: a letter and questionnaire were given to patients to be completed by their GP or nurse (district or practice), who was requested to take wound swab if a SWI was suspected
- Patients were contacted by telephone at 1 month postoperatively.
- Researchers contacted the GPs and patients’ notes were inspected by an infection control nurse at the outpatient clinic.
### TABLE 23 contd Prospective studies of surgical wound infection

<table>
<thead>
<tr>
<th>Study</th>
<th>Study design, sample size, aim</th>
<th>Definition, grading system, cited references</th>
<th>Data collection, postdischarge surveillance</th>
</tr>
</thead>
<tbody>
<tr>
<td>Jewesson et al., 1997.94 Canada</td>
<td>RCT of antibiotic prophylaxis in elective colorectal surgery (n = 153)</td>
<td><strong>Definition</strong>: used ASEPSIS and a classification scheme for the definition of wound infection. <strong>Reference to definition</strong>: Wilson et al.90</td>
<td><strong>Hospital</strong>: wounds were assessed daily, by house and investigative personnel, for appearance and proportion of wound infection. Any additional treatment (e.g., drainage, debridement, antibiotics, prolonged stay) were scored. The sum of the daily scores for 5 of the first 7 postoperative days and the additional treatment scores were summed to give the total ASEPSIS score for each patient. <strong>Community</strong>: contacted by telephone 30 days after surgery to assess the status of the wounds. If infection was suspected, the surgeon was contacted to determine the nature of the infection and the treatment required.</td>
</tr>
</tbody>
</table>

| Kingston et al., 1995.107 UK | Prospective study of elective colorectal surgery (n = 618) | **Definition**: purulent discharge from a wound (abdominal or perineal), a major wound infection was defined as a purulent discharge associated with pain and/or pyrexia and positive bacteriology. **Reference to definition**: none | **Hospital**: wounds were inspected daily and observations recorded. All patients were supervised by clinical research nurses until discharge. **Community**: at the time of discharge patients were requested to inform the nurse of any septic or non-septic complications after discharge. |

| Kluytmans et al., 1994.41 Netherlands | Prospective surveillance study in thoracic surgery (n = 983) | **Definition**: 1988 CDC31 | **Hospital**: patients were visited by an investigator and their hospital records, notes and reports were assessed for signs of infection. **Community**: no details given on follow-up after discharge. |

| Kotisso and Aseffa, 1998.84 Ethiopia | Prospective study of abdominal surgery (n = 129) | **Definition**: wounds were inspected for clinical signs of infection, such as erythema, oedema and discharge. **Reference to definition**: none | **Hospital**: daily inspection of wound, until discharge, for clinical signs of infection such as erythema, oedema and discharge. **Community**: patients staying for more than 10 days because of wound infection were considered to have delayed infection. |

| Kow et al., 1995.76 Australia | RCT of antibiotic prophylaxis in abdominal surgery (n = 1010) | **Definition**: the presence of purulent discharge from the wound or a serous discharge with a positive culture of pathogenic organism(s). **Reference to definition**: none | **Hospital**: wounds were assessed daily, by a clinical trial monitor, for evidence of infection. Swabs were obtained from patients with evidence of discharge from their surgical wounds. **Community**: wounds were reviewed in the outpatient department during weeks 4 to 6 postoperatively for evidence of wound infections. The final assessment was completed at this time. |

**Notes**: RCT, randomised controlled trial; SWI, surgical wound infection.
<table>
<thead>
<tr>
<th>Study</th>
<th>Study design, sample size, aim</th>
<th>Definition, grading system, cited references</th>
<th>Data collection, postdischarge surveillance</th>
</tr>
</thead>
<tbody>
<tr>
<td>Kumara-krishnan et al., 1997, India</td>
<td>RCT of antibiotic prophylaxis in appendicectomy (n = 152)</td>
<td><strong>Definition</strong>: declared as infected or non-infected following the guidelines for SSIs (i.e. superficial/deep SSIs)</td>
<td><strong>Hospital</strong>: wounds were evaluated by a single observer to eliminate observer bias. Postoperatively, wounds were inspected daily for signs of infection, until discharge (with a minimum of at least 4 days in hospital)</td>
</tr>
<tr>
<td></td>
<td><strong>Aim</strong>: to study the efficacy of three regimens of antimicrobial drug combinations in reducing postoperative wound sepsis in acute appendicitis</td>
<td><strong>Reference to definition</strong>: Sheridan and co-workers42</td>
<td><strong>Community</strong>: on discharge patients were asked to report immediately if any pain, fever or discharge from wound was noted, and to attend for review on day 30. In 10/21 (48%) patients infections developed after discharge; 24/152 (16%) failed to complete the follow-up</td>
</tr>
<tr>
<td></td>
<td><strong>Grading system</strong>: none</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Kurz et al., 1996, USA, and the Austrian Study of Wound Infection and Temperature Group</td>
<td>RCT of intraoperative warming in colorectal surgery (n = 200)</td>
<td><strong>Definition</strong>: the presence of pus and a positive culture</td>
<td><strong>Hospital</strong>: evaluated daily by a blinded physician. Wounds were suspected of being infected when pus could be expressed from the surgical incision or be aspirated from a localised mass inside the wound. All wound infections diagnosed within 15 days were included</td>
</tr>
<tr>
<td></td>
<td><strong>Aim</strong>: to test the hypothesis that mild core hyperthermia increases both the incidence of SWI and the length of hospitalisation</td>
<td><strong>Reference to definition</strong>: none</td>
<td><strong>Community</strong>: blinded-physician assessment at 2 weeks</td>
</tr>
<tr>
<td></td>
<td><strong>Grading system</strong>: ASEPSIS135</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Kurz et al., 1996, Belgium</td>
<td>Analysis of antimicrobial prophylaxis data from a prospective surveillance programme</td>
<td><strong>Definition</strong>: 1988 CDC11</td>
<td><strong>Hospital</strong>: data were collected as part of the NSIH in Belgium. Baseline data and data on patients with SWI were collected for the duration of hospital stay by individual hospitals</td>
</tr>
<tr>
<td></td>
<td><strong>Aim</strong>: to evaluate the current practice of surgical antimicrobial prophylaxis in Belgium</td>
<td><strong>Reference to definition</strong>: Garner and co-workers11</td>
<td><strong>Community</strong>: data were collected by some participating hospitals, but no details were given</td>
</tr>
<tr>
<td></td>
<td><strong>Grading system</strong>: none</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Lecuona et al., 1998, Spain</td>
<td>Prospective surveillance in general surgery (n = 1103)</td>
<td><strong>Definition</strong>: 1992 CDC12</td>
<td><strong>Hospital</strong>: no details were given about wound assessment, but patients with nosocomial infections were prospectively identified according to the CDC criteria</td>
</tr>
<tr>
<td></td>
<td><strong>Aim</strong>: to analyse the risk factors associated with SSI diagnosed postdischarge</td>
<td><strong>Reference to definition</strong>: Horan and co-workers13</td>
<td><strong>Community</strong>: surveillance was extended to 30 days postdischarge based on the scheduled visit to the surgeon. Follow-up was complete for 70.4% of patients</td>
</tr>
<tr>
<td></td>
<td><strong>Grading system</strong>: none</td>
<td></td>
<td></td>
</tr>
<tr>
<td>L’Ecuyer et al., 1996, USA</td>
<td>Prospective surveillance study in CABG and valve surgery (n = 1554)</td>
<td><strong>Definition</strong>: 1992 CDC definition.32</td>
<td><strong>Hospital</strong>: wound infections were confirmed by direct bedside examination and review of the patient’s medical records by an infection control nurse, using the NNIS (CDC) definitions</td>
</tr>
<tr>
<td></td>
<td><strong>Aim</strong>: to describe the clinical nature and outcome of all chest and leg infections following CABG, valve and combined CABG + valve surgery</td>
<td><strong>Deep incisional and deep organ/space infections were considered together because of the difficulty in distinguishing between the two in the chest</strong></td>
<td><strong>Community</strong>: no formal postdischarge surveillance was performed</td>
</tr>
<tr>
<td></td>
<td><strong>Reference to definition</strong>: Sawyer and Pruett43</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td><strong>Grading system</strong>: none</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

**RCT**, randomised controlled trial; **SSI**, surgical site infection; **SWI**, surgical wound infection

*continued*
## TABLE 23 contd  Prospective studies of surgical wound infection

<table>
<thead>
<tr>
<th>Study</th>
<th>Study design, sample size, aim</th>
<th>Definition, grading system, cited references</th>
<th>Data collection, postdischarge surveillance</th>
</tr>
</thead>
<tbody>
<tr>
<td>Liberman et al., 1995, USA</td>
<td>RCT of antibiotic prophylaxis in non-perforated appendicectomy (n = 136)</td>
<td>Definition: criteria based on the clinical appearance of the wound. If peri-incisional erythema and incisional drainage were present, the wound was classified as infected. Reference to definition: none</td>
<td>Hospital: all patients were examined daily throughout their hospital stay. Community: patients were followed up at 3 weeks postoperatively and asked to attend the clinic if complications developed before this time.</td>
</tr>
<tr>
<td></td>
<td><strong>Aim</strong>: to compare single-dose prophylaxis with cefotaxin with single-dose cefoxitin</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td><strong>Definition</strong>: serious wound infection was defined as the presence of pus or sanguinopurulent discharge at the operative site. Reference to definition: none</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td><strong>Grading system</strong>: none</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Liem et al., 1997, Netherlands</td>
<td>RCT of open versus laparoscopic inguinal herniorrhaphy (n = 1051)</td>
<td></td>
<td>Hospital: data collection was performed by the attending resident or surgeon, and each patient was evaluated at the hospital monthly. Community: patients were requested to return at 1 and 6 weeks, 6 months, and 1 and 2 years. Most cases were assessed by the surgeon who had performed the surgery. All patients were visited or contacted by telephone by a blinded member of the study soon after discharge to stress the importance of follow-up. Home visits by experience physicians were also conducted at 1 and 2 years postoperatively if patients did not want to, or were unable to, attend hospital.</td>
</tr>
<tr>
<td></td>
<td><strong>Aim</strong>: to compare conventional anterior repair with extraperitoneal laparoscopic repair in terms of postoperative recovery, complications and recurrence rates in patients with primary or first recurring unilateral hernia</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td><strong>Reference to definition</strong>: none</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td><strong>Grading system</strong>: none</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Lizan-Garcia et al., 1997, Spain</td>
<td>Prospective study in general surgery (n = 2237)</td>
<td>Definition: 1988 CDC31</td>
<td>Hospital: information was collected every 2 days by nurses trained in epidemiology. Only wound infections that occurred by the time of discharge were included in the study. The mean length of hospital stay was 16 days (standard deviation 14). Community: no follow-up.</td>
</tr>
<tr>
<td></td>
<td><strong>Aim</strong>: to quantify the surgical infection rate, to assess adherence with the antibiotic prophylaxis protocol, and to identify independent factors associated with SWI</td>
<td>Reference to definition: Garner and co-workers1</td>
<td></td>
</tr>
<tr>
<td></td>
<td><strong>Reference to definition</strong>: none</td>
<td>Grading system: none</td>
<td></td>
</tr>
<tr>
<td></td>
<td><strong>Definition</strong>: a wound having at least two of the three criteria reported by a patient: (1) one of the wound complications listed under ‘Community’ in the next column; and/or (2) antibiotic treatment for a SWI; and/or (3) physician diagnosis</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td><strong>Reference to definition</strong>: none</td>
<td><strong>Grading system</strong>: none</td>
<td></td>
</tr>
<tr>
<td>Manian and Meyer, 1993, USA</td>
<td>Prospective surveillance of inpatient and outpatient surgical procedures (n = 501, randomly selected from a population of 7433).</td>
<td></td>
<td>Hospital: no details were given on hospital assessment as the aim was to evaluate the postdischarge surveillance method. Community: patient telephone surveys were undertaken over 4 separate months to contact 501 patients at approximately day 30 postoperatively. A standardised questionnaire was used to record: (a) presence of pus or yellowish discharge; (b) persistence of pain or redness around the incision; (c) poor wound healing; and (d) persistent or intermittent fever. Patients were also asked about follow-up visits, antibiotic treatment, re-admission, and whether the surgeon ever diagnosed an SWI. The follow-up contact by telephone was only achieved in 38% (n = 189) of cases. (See section on postdischarge surveillance, p. 34)</td>
</tr>
<tr>
<td></td>
<td><strong>Aim</strong>: to determine whether concomitant patient telephone surveys would provide additional useful information regarding the status of surgical wounds in the outpatient setting</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

**RCT**, randomised controlled trial; **SWI**, surgical wound infection

**continued**
### TABLE 23 contd Prospective studies of surgical wound infection

<table>
<thead>
<tr>
<th>Study</th>
<th>Study design, sample size, aim</th>
<th>Definition, grading system, cited references</th>
<th>Data collection, postdischarge surveillance</th>
</tr>
</thead>
<tbody>
<tr>
<td>Manian et al., 1997, 1998, USA (Same data in both studies, but a subset was used in the second survey)</td>
<td>Analysis of prospectively collected data in ambulatory surgery (n = 13 specialties, n = 156,977 procedures)</td>
<td><strong>Definition:</strong> 1988 CDC, although the authors refer to 'SSI' and thus it should be 1992 CDC <strong>Reference to definition:</strong> Garner</td>
<td><strong>Hospital:</strong> surveillance was by infection control staff and diagnosis by patient chart review, staff consultation and routine examination of suspicious wounds. Microbiology laboratory results were perused daily. For ambulatory surgery patients, the surgeon was contacted and asked to provide details of the site and presence of possible infection <strong>Community:</strong> monthly computer-generated questionnaires were sent to each surgeon, to be returned to the department of infection control. Each questionnaire contained the CDC definition, a list of operated cases and options for infection (SSI (yes/no/no follow-up), date of SSI diagnosis, any culture results). The response rate by surgeons from different specialties ranged from 52% to 96%</td>
</tr>
<tr>
<td>Matikainen and Hiltunen, 1993, Finland</td>
<td>RCT of antibiotic prophylaxis in colorectal surgery (n = 628)</td>
<td><strong>Definition:</strong> wound sepsis (suppuration of the wound or positive bacterial culture of the wound)</td>
<td><strong>Hospital:</strong> wounds were inspected daily by the surgeon <strong>Community:</strong> all patients were assessed by the surgeon at an outpatient follow-up visit on average 4 weeks postoperatively. Patients were instructed to contact the hospital at any time if they suspected a wound complication</td>
</tr>
<tr>
<td>Medina-Cuadros et al., 1996, Spain</td>
<td>Prospective study of general surgical patients (n = 1483)</td>
<td><strong>Definition:</strong> 1988 CDC <strong>Reference to definition:</strong> Garner</td>
<td><strong>Hospital:</strong> two of the authors visited patients on a daily basis to collect data and observe wounds <strong>Community:</strong> surveillance was extended to 30 days postdischarge. All emergency department forms were reviewed and questionnaires were sent monthly to surgeons to complete information on SWIs detected in discharged patients. 90% of surgeons responded to every monthly questionnaire for the 18-month study. Passive surveillance was based on the assumption that the patient will present if a problem arises</td>
</tr>
<tr>
<td>Medina et al., 1997, Spain</td>
<td>Prospective observational study of abdominal herniorrhaphy (n = 497)</td>
<td><strong>Definition:</strong> 1988 CDC <strong>Reference to definition:</strong> Garner and co-workers</td>
<td><strong>Hospital:</strong> two authors assessed patients daily <strong>Community:</strong> surveillance was extended to 30 days after discharge. Patients were actively sought among emergency department admission forms. A monthly questionnaire was sent to other surgeons in the department in which information was asked for about SWIs that had been detected in discharged patients</td>
</tr>
</tbody>
</table>

*RCT, randomised controlled trial; SSI, surgical site infection; SWI, surgical wound infection*
<table>
<thead>
<tr>
<th>Study</th>
<th>Study design, sample size, aim</th>
<th>Definition, grading system, cited references</th>
<th>Data collection, postdischarge surveillance</th>
</tr>
</thead>
<tbody>
<tr>
<td>Milsom et al., 1998.** USA</td>
<td>RCT of antibiotic prophylaxis in colorectal surgery (n = 518)</td>
<td><strong>Definition</strong>: as per the grading system below</td>
<td><strong>Hospital</strong>: daily assessment of surgical sites by investigator or by a designate</td>
</tr>
<tr>
<td></td>
<td><strong>Aim</strong>: to compare the prophylactic efficacy and safety of a single intravenous dose of alatrofloxacin with intravenous cefotetan in the prevention of postoperative wound infections</td>
<td><strong>Reference to definition</strong>: none</td>
<td><strong>Community</strong>: 30 days after surgery, wounds were examined for signs or symptoms of a primary infection. 24 patients were lost to follow-up</td>
</tr>
<tr>
<td></td>
<td></td>
<td><strong>Grading system</strong>: wound infection graded as:</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td><strong>Minor</strong>: erythema extending at least 2 cm from the wound in any direction</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td><strong>Major</strong>: wound infection with erythema and drainage, with purulent drainage, or a wound that had been opened and not re-closed</td>
<td></td>
</tr>
<tr>
<td>Mishriki et al., 1993,** UK</td>
<td>Prospective study of elective general surgery (n = 702)</td>
<td><strong>Definition</strong>: salient features of wound (erythema, cellulitis, purulent/non-purulent discharge or breakdown) were recorded</td>
<td><strong>Hospital</strong>: wounds were routinely examined by medical staff for signs of infection while patients were in hospital and before discharge</td>
</tr>
<tr>
<td></td>
<td><strong>Aim</strong>: to review the apparent variation in the wound infection rate according to the application of different criteria of wound infection</td>
<td><strong>Reference to definition</strong>: none</td>
<td><strong>Community</strong>: at discharge patients were given a full explanation and written instructions to seek advice if symptoms developed, and were given a letter/questionnaire to pass to their consulting doctor. A random postal survey of 80 patients was undertaken to determine the accuracy of wound infection reporting. Wound infection rates were presented according to individual criteria (e.g. cellulitis only)</td>
</tr>
<tr>
<td>Mitchell et al., 1999,** Australia</td>
<td>Prospective study of multiple elective major procedures (n = 1360)</td>
<td><strong>Definition</strong>: 1992 CDC**</td>
<td><strong>Hospital</strong>: a research nurse collected data on operations and clinical assessment performed on day 5 postoperatively or later</td>
</tr>
<tr>
<td></td>
<td><strong>Aim</strong>: to evaluate two methods of postdischarge surgical wound surveillance to compare the incidence and outcomes of wound infection that develop prior to patients' discharge with those that develop after discharge</td>
<td><strong>Reference to definition</strong>: Horan and co-workers**</td>
<td><strong>Community</strong>: each patient was given a questionnaire and a mail-back form to return after day 28 postoperatively. Surgeons also received a mail-back form to complete on wound status. If a returned form indicated a SWI, the patient and surgeon were contacted by the research nurse for further details and diagnostic confirmation. Agreement values were given between patients and surgeons, with ( \kappa = 0.73 ), although more patients reported SWIs than did surgeons</td>
</tr>
<tr>
<td></td>
<td></td>
<td><strong>Grading system</strong>: none</td>
<td></td>
</tr>
<tr>
<td>Moro et al., 1996,** Italy, PRINOS Study Group</td>
<td>Prospective multicentre study in general and thoracic surgery (n = 2263)</td>
<td><strong>Definition</strong>: wounds were considered infected if discharge of pus was present. Wounds with serous or non-purulent discharge and positive cultures were considered infected only if significant physical signs were concurrently present</td>
<td><strong>Hospital</strong>: patients were visited daily by the intern surgeon to detect infections</td>
</tr>
<tr>
<td></td>
<td><strong>Aim</strong>: to investigate the potential risk factors for SWI among hospital patients who underwent clean operations</td>
<td><strong>Reference to definition</strong>: none</td>
<td><strong>Community</strong>: infections after discharge were not actively surveyed</td>
</tr>
<tr>
<td></td>
<td></td>
<td><strong>Grading system</strong>: none</td>
<td></td>
</tr>
</tbody>
</table>

**RCT, randomised controlled trial; SWI, surgical wound infection**

continued
### TABLE 23 contd  Prospective studies of surgical wound infection

<table>
<thead>
<tr>
<th>Study</th>
<th>Study design, sample size, aim</th>
<th>Definition, grading system, cited references</th>
<th>Data collection, postdischarge surveillance</th>
</tr>
</thead>
<tbody>
<tr>
<td>Nichols et al., 1993, USA</td>
<td>Double-blind RCT of antibiotic therapy for penetrating abdominal trauma (n = 170)</td>
<td><strong>Definition</strong>: as per the grading system below</td>
<td><strong>Hospital</strong>: while in hospital patients were monitored daily for the development of infection and adverse effects</td>
</tr>
<tr>
<td></td>
<td><strong>Aim</strong>: to investigate changes in antibiotic duration while comparing two commonly used antimicrobial regimens and selective incisional wound closure</td>
<td><strong>Reference to definition</strong>: none</td>
<td><strong>Community</strong>: followed-up at the outpatient clinic up to 30 days postdischarge</td>
</tr>
<tr>
<td></td>
<td></td>
<td><strong>Grading system</strong>: No infection or incidental infections: minor wound infections (e.g. incisional or stitch abscesses) that did not require additional antibiotic therapy or intervention beyond normal wound care were included in this category</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>Minor infection: infections were placed in this category if they were (1) nosocomial (urinary or respiratory tract) and were unrelated to the trauma, or (2) were localised to the surgical incision and required surgical intervention or a change in the antibiotic therapy</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>Major infection: infections were included in this group if they were related to the original trauma or surgery and if surgical intervention or a change in antibiotics was required. This category included intra-abdominal infections and sepsicaemia</td>
<td></td>
</tr>
<tr>
<td>Noel et al., 1997, UK</td>
<td>Prospective study of clean elective surgery (n = 166)</td>
<td><strong>Definition</strong>: infection taken to be redness and/or discharge as noted on the questionnaire</td>
<td><strong>Hospital</strong>: no details given</td>
</tr>
<tr>
<td></td>
<td><strong>Aim</strong>: to measure the incidence of postoperative wound infection after clean surgery in the 4 weeks following early discharge and its effect on community medical services</td>
<td><strong>Reference to definition</strong>: questionnaire based on a pilot study[44]</td>
<td><strong>Community</strong>: a questionnaire survey, asking about discomfort, pain, infection (redness, discharge), to patients after discharge. Two questionnaires were used in the study, one for patients and one for GPs and practice/district nurses. 11/12 patients who reported infection had this validated by a doctor or nurse</td>
</tr>
<tr>
<td>Oertli et al., 1994, Switzerland</td>
<td>RCT of tranexamic acid versus placebo in lumpectomy or mastectomy (n = 160)</td>
<td><strong>Definition</strong>: infection was defined as the emergence of pus from the incision wound or suction drain, with positive bacterial culture</td>
<td><strong>Hospital</strong>: not described</td>
</tr>
<tr>
<td></td>
<td><strong>Aim</strong>: not clearly stated</td>
<td><strong>Reference to definition</strong>: none</td>
<td><strong>Community</strong>: final check for bruising, haematoma and seroma was on day 14 when the sutures were removed by the GP or surgeon</td>
</tr>
<tr>
<td>Palmer et al., 1994, UK</td>
<td>RCT of antibiotic prophylaxis in abdominal surgery (n = 509)</td>
<td><strong>Definition</strong>: erythema, discharge or dehiscence</td>
<td><strong>Hospital</strong>: assessed by a designated research nurse, blinded, at 72 hours, and then daily</td>
</tr>
<tr>
<td></td>
<td><strong>Aim</strong>: to compare the safety and clinical efficacy of co-amoxiclav with cefuroxime plus metronidazole in patients undergoing abdominal surgery</td>
<td><strong>Reference to definition</strong>: none</td>
<td><strong>Community</strong>: reviewed for infection, by the research nurse 21–35 days postoperatively</td>
</tr>
</tbody>
</table>

**RCT**, randomised controlled trial
### TABLE 23 contd  Prospective studies of surgical wound infection

<table>
<thead>
<tr>
<th>Study</th>
<th>Study design, sample size, aim</th>
<th>Definition, grading system, cited references</th>
<th>Data collection, postdischarge surveillance</th>
</tr>
</thead>
</table>
| Platell and Hall, 1997, Australia | Study using prospective data from three clinical trials of colorectal and abdominal procedures (n = 553) | **Definition:** either a purulent discharge or a serous discharge with the culture of pathogenic organisms  
**Reference to definition:** none  
**Grading system:** none | **Hospital:** not specified  
**Community:** not specified, although the results suggest that one-third of wound infections presented after hospital discharge |
| Poulsen et al., 1995, Denmark | Cohort study with retrospective analysis from multiple surgical procedures (n = 4515) | **Definition:** superficial wound infection confined to the incision; deep wound infection involves tissues adjacent to the wound, such as subfacial layers and intra-abdominal structures  
**Reference to definition:** Simmons  
**Grading system:** none | **Hospital:** not given  
**Community:** not given |
| Rantala et al., 1997, Finland | Prospective study of multiple general surgical procedures (n = 807) | **Definition:** 1992 CDC  
**Reference to definition:** Horan and co-workers  
**Grading system:** none | **Hospital:** wound surveillance was performed on each patient by a research nurse during hospital stay  
**Community:** all patients were advised to monitor their wounds postdischarge and asked to contact personnel if an abnormality was suspected. All were given written instructions to inform personnel taking care of the wounds postdischarge, including instruction to the attending physician to obtain culture from the wound discharge. All patients were contacted by mail or telephone at 1 month postoperatively and asked about healing and contact with medical personnel. Final follow-up (n = 807; n = 35 excluded) was by telephone (73.8%), mailed questionnaire (15.8%) or follow-up visit (7.2%). The authors reported that telephone contact was laborious and 98.5% follow-up was achieved by combining telephone contact and a mailed questionnaire |
| Reggiori et al., 1996, Uganda | Prospective study of hysteroscopy, ectopic pregnancy surgery, hysterectomy and caesarean section (n = 850) | **Definition:** as per the grading system below  
**Reference to definition:** none  
**Grading system:** grade 1, superficial infection (cellulitis with minimal purulent exudate); grade 2, deep infection (cellulitis with moderate purulent exudate); grade 3, infection throughout the wound, with or without dehiscence  
**Reference to grading system:** Karl and co-workers | **Hospital:** assessed each day by two supervisors aware of the type of prophylaxis used  
**Community:** follow-up performed 2 weeks postoperatively, with cash incentives |
TABLE 23 contd  Prospective studies of surgical wound infection

<table>
<thead>
<tr>
<th>Study</th>
<th>Study design, sample size, aim</th>
<th>Definition, grading system, cited references</th>
<th>Data collection, postdischarge surveillance</th>
</tr>
</thead>
</table>
| Renz and Feliciano, 1995, USA | Prospective observational study of unnecessary laparotomies (no repairs/no drains) | Definition: 1988 CDC<sup>11</sup>  
Reference to definition: Garner and co-workers<sup>11</sup> | Hospital: examined twice daily by the primary author  
Community: not recorded |
|                               | Aim: a prospective study to record all perioperative complications | Grading system: none | |
| Roberts et al., 1998, Canada  | Prospective study of the influence of surveillance methods on the spinal/trauma surgery unit | Definition: 1992 CDC<sup>12</sup>  
Reference to definition: Horan and co-workers<sup>12</sup> | Hospital: surveillance performed using a continually updated list of procedures from operating room and nursing units; visited by the infection control nurse twice weekly for chart review and staff consultation  
Community: surgeons were asked to complete and return an infection survey form for surgery performed on patients in the previous 6 weeks. At 6-week intervals, they were asked to complete and return new forms. Asked to mark yes/no for infection, with space given for details of infection. 100% compliance from surgeons |
|                               | Aim: to assess the role played by the use of postdischarge surveillance and the pattern of surgical procedures performed | Grading system: none | |
| Saha, 1996, UK                | Prospective study of abdominal surgery for intra-abdominal sepsis (n = 182) | Definition: evidence of organisms present in wound discharge  
Reference to definition: none | Hospital: no details given  
Community: reviewed at 1 month |
|                               | Aim: to evaluate the efficacy of metronidazole lavage in the treatment of intra-peritoneal sepsis | Grading system: none | |
| Salem et al., 1994, UAE       | RCT of antibiotic prophylaxis in appendicectomy (n = 330) | Definition: discharge of pus or positive bacteriological culture from a wound discharge. A stitch abscess, remote from the incision, or erythema that did not progress to suppuration were excluded  
Reference to definition: Ljungqvist<sup>121</sup> | Hospital: incisions were not inspected routinely  
Community: reviewed at outpatients department on day 14 postoperatively and 6 weeks later |
|                               | Aim: to compare cefoxitin with piperacillin for the prevention of postoperative wound infection in patients with acute non-perforated appendicitis | Grading system: none | |
| Santos et al., 1997, Brazil   | Prospective surveillance study in general surgery | Definition: 1992 CDC<sup>12</sup>  
Reference to definition: Horan and co-workers<sup>12</sup> | Hospital: wards visited daily by an investigator and wounds examined for the presence of infection  
Community: patients examined at the outpatient surgery clinic within 30 days of surgery. Patients were instructed to return a postal questionnaire if they experienced wound problems. The majority (87.5%) of infections occurred postdischarge |
|                               | Aim: to evaluate the importance and feasibility of postdischarge surveillance of the wound infection rate in herniorrhaphy at the University Hospital of the Federal University of Rio de Janeiro | Grading system: none | |

<sup>RCT, randomised controlled trial</sup> continued
### TABLE 23 contd Prospective studies of surgical wound infection

<table>
<thead>
<tr>
<th>Study</th>
<th>Study design, sample size, aim</th>
<th>Definition, grading system, cited references</th>
<th>Data collection, postdischarge surveillance</th>
</tr>
</thead>
<tbody>
<tr>
<td>Sayed and Cade, 1996, Australia</td>
<td>Non-randomised comparative study of appendicectomy (n = 92)</td>
<td>Definition: definite wound infection if swab culture confirmed an organism in the presence of clinical infection. Possible wound infection when no organism was isolated in the presence of ooz or redness</td>
<td>Hospital: no details given Community: no details given</td>
</tr>
<tr>
<td>Schein et al., 1994, Israel</td>
<td>Prospective non-randomised study of laparotomy for suspected intra-abdominal infection (n = 163)</td>
<td>Definition: wound purulence requiring early removal of sutures and drainage</td>
<td>Hospital: no details given Community: no details given</td>
</tr>
<tr>
<td>Serour et al., 1996, Israel</td>
<td>Prospective study of appendicectomy in children (n = 216)</td>
<td>Definition: redness, oedema, swelling and discharge of pus</td>
<td>Hospital: wounds examined for redness, oedema, swelling and pus Community: reviewed at outpatient clinic at 1 week and 1 month after discharge. No details given on losses to follow-up</td>
</tr>
<tr>
<td>Shirahatti et al., 1993, India</td>
<td>RCT of skin preparation in elective and emergency general surgery (n = 135)</td>
<td>Definition: infected if wound showed redness or swelling of the surrounding area or had a discharge, irrespective of whether any organisms were grown in the discharge</td>
<td>Hospital: no details given Community: patients with clean procedures and early discharge were followed up the next day in the outpatient clinic</td>
</tr>
<tr>
<td>Siegman-Igra et al., 1993, Israel</td>
<td>Prospective study of gastrointestinal surgery (n = 813)</td>
<td>Definition: either the clinical observation of pus in the wound, or of discharge other than pus, provided that two of the following stipulations are met: repeated growth of same organism in culture, systemic treatment with antibiotics, or local treatment such as draining</td>
<td>Hospital: trained nurse epidemiologists followed operated patients daily until discharge and visited their bedside at least three times weekly to observe wounds. They also used case notes and staff consultation. They were closely supervised by two rotating central team nurses, who evaluated the quality of the data. Reported accuracy rates in pre-1993 publications Community: postdischarge infections estimated by repeated telephone contacts with a 20% subsample of patients. However, low proportions of SWI were found following gastrointestinal surgery and therefore this study used hospital-reported infection rates</td>
</tr>
</tbody>
</table>

*RCT, randomised controlled trial; SWI, surgical wound infection*
### TABLE 23 contd  Prospective studies of surgical wound infection

<table>
<thead>
<tr>
<th>Study</th>
<th>Study design, sample size, aim</th>
<th>Definition, grading system, cited references</th>
<th>Data collection, postdischarge surveillance</th>
</tr>
</thead>
<tbody>
<tr>
<td>Simchen et al., 1996, Israel</td>
<td>Analysis of data from a multicentre (11 hospitals) prospective study of different surgical procedures (n = 5571 patients)</td>
<td><strong>Definition</strong>: either the presence of pus in the wound or the presence of discharge other than pus, if two of the following conditions are met: repeated culture with the same organism, use of systemic antibiotics, or local treatment, such as evacuation or draining</td>
<td><strong>Hospital</strong>: trained nurse epidemiologists followed operated patients daily until discharge and visited their bedside at least three times weekly to observe wounds. They also used case notes and staff consultation</td>
</tr>
<tr>
<td></td>
<td><strong>Aim</strong>: to develop a method for analysing differences in the performance of hospitals with respect to outcome by separating patient factors from procedural factors</td>
<td><strong>Reference to definition</strong>: none</td>
<td><strong>Community</strong>: for 80% of patients, follow-up terminated on the day of discharge. Home follow-up by telephone interview was conducted with the remainder of patients</td>
</tr>
<tr>
<td>Slaughter et al., 1993, USA</td>
<td>Analysis of prospectively collected wound infection data in cardiac surgery (n = 2405)</td>
<td><strong>Definition</strong>: infected if purulent material discharged from the wound, with or without a positive culture, or the responsible surgeon deemed the wound infected based on clinical judgement. For chest wounds, a superficial infection included those infections limited to the subcutaneous tissues. A major chest wound infection included all cases where tissues were opened down to the sternal wires or beyond (including mediastinitis)</td>
<td><strong>Hospital</strong>: prospective surveillance of wounds by nurse epidemiologist by: (a) daily visits for direct inspection of suspicious wounds; (b) daily review of postoperative wound cultures at the microbiology laboratory; and (c) continuous contact by the nurse epidemiologist with the ward and clinic nurses. In no case could the surgeon overrule the diagnosis of infection by the nurse epidemiologist</td>
</tr>
<tr>
<td></td>
<td><strong>Aim</strong>: to review the results of the ongoing wound surveillance programme after coronary artery bypass operations to look for changing trends in wound infections that might be adjusted to improve patient care</td>
<td><strong>Reference to definition</strong>: none</td>
<td><strong>Community</strong>: all patients were seen at the clinic at 4 weeks post-operatively. Therefore, all wounds were subject to postdischarge surveillance with a minimum of 30 days follow-up</td>
</tr>
<tr>
<td>Smack et al., 1996, USA</td>
<td>RCT of wound care in biopsies and Mohs surgery in an outpatient dermatology clinic (n = 922)</td>
<td><strong>Definition</strong>: infection defined by the presence of three symptoms (pus, erythema, tenderness) and a positive culture demonstrating a pathogenic bacterial strain</td>
<td><strong>Hospital</strong>: not relevant, as the study was conducted in an outpatient dermatology clinic</td>
</tr>
<tr>
<td></td>
<td><strong>Aim</strong>: to assess the effect of white petroleum versus bacitracin ointment on wound infection and allergy incidence</td>
<td><strong>Reference to definition</strong>: none</td>
<td><strong>Community</strong>: the duration of participation was 4 weeks. Patients were educated on the signs and symptoms (tenderness, redness, purulence) of infection and instructed to observe and dress the wound for 7–10 days. Patients with punch biopsies attended the clinic on days 7 and 28.Mohs/dermabrasions were followed up at days 1, 7 and 28. All patients were asked to return if there was any sign of infection. All patients unable to return for follow-up were contacted by telephone after the expiration of study and questioned about the post-operative course of their wound</td>
</tr>
<tr>
<td></td>
<td><strong>Grading system</strong>: clinical parameters of presence/absence of pus, erythema, tenderness, itch, graded on a subjective scale where: (–) not present, (+) minimally present, (++) extensively present</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Stahle et al., 1997, Sweden</td>
<td>Prospective study of CABG and valve surgery (n = 13,285)</td>
<td><strong>Definition</strong>: sternal wound complication was classified as any sternal wound complication requiring re-operation with the aim of sternal re-stabilisation and/or drainage necessitated by instability of the sternal wound, suspected mediastinitis or other signs of deep wound infection</td>
<td><strong>Hospital</strong>: not specified</td>
</tr>
<tr>
<td></td>
<td><strong>Aim</strong>: to investigate the incidence of sternal wound complication and associated mortality, infection and to identify and evaluate risk factors for this evaluation</td>
<td><strong>Reference to definition</strong>: none</td>
<td><strong>Community</strong>: not given, but the authors state that late infections up to month 7 postoperatively were included</td>
</tr>
<tr>
<td></td>
<td><strong>Grading system</strong>: none</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

**RCT**: randomised controlled trial
<table>
<thead>
<tr>
<th>Study</th>
<th>Study design, sample size, aim</th>
<th>Definition, grading system, cited references</th>
<th>Data collection, postdischarge surveillance</th>
</tr>
</thead>
<tbody>
<tr>
<td>Stewart et al., 1995.</td>
<td>Multicentre RCT of antibiotic prophylaxis in colorectal surgery (n = 379)</td>
<td><strong>Definition:</strong> wound infections occurring within 42 days of operation were considered to be evidence of failure of antibiotic prophylaxis. Diagnoses were made on clinical criteria, with microbiological confirmation whenever possible.</td>
<td>Hospital: not specified</td>
</tr>
<tr>
<td></td>
<td><strong>Aim:</strong> to assess the efficacy of piperacillin by adding sulbactam to the prophylactic regimen</td>
<td><strong>Reference to definition:</strong> Peel et al. 29</td>
<td>Community: all patients were reviewed at 6 weeks. No details were given of how or by whom.</td>
</tr>
<tr>
<td></td>
<td></td>
<td><strong>Grading system:</strong> none</td>
<td></td>
</tr>
<tr>
<td>Sturgis et al., 1996.</td>
<td>RCT of antibiotic prophylaxis in percutaneous endoscopic gastrostomy (n = 115)</td>
<td><strong>Definition:</strong> wound infection if pus present or the maximum combined score was greater than 8</td>
<td>Hospital: the gastrostomy tube site was evaluated daily for 7 days by an investigator. The peristomal area was assessed daily for erythema, induration and exudate.</td>
</tr>
<tr>
<td></td>
<td><strong>Aim:</strong> to determine whether prophylactic cefazolin before percutaneous endoscopic gastrostomy reduces or prevents the incidence of peristomal wound infections</td>
<td><strong>Reference to definition:</strong> Jain and co-workers 15</td>
<td>Community: patients discharged to a nursing facility before the end of observation week (n = 13) were followed up by telephone by a nurse. If infection was suspected, the patient was evaluated by an investigator.</td>
</tr>
<tr>
<td></td>
<td></td>
<td><strong>Grading system:</strong> none</td>
<td></td>
</tr>
<tr>
<td>Taylor et al., 1997.</td>
<td>RCT of antibiotic prophylaxis of hernia repair surgery (n = 619)</td>
<td><strong>Definition:</strong> clinical criteria, purulent wound discharge or spreading erythema indicative of cellulitis, wound breakdown with clinical evidence of infection.</td>
<td>Hospital: no details given</td>
</tr>
<tr>
<td></td>
<td><strong>Aim:</strong> to compare co-amoxiclav with placebo in patients undergoing open groin hernia repair</td>
<td><strong>Reference to definition:</strong> Peel et al. 29</td>
<td>Community: diary cards were given to patients and a nurse or GP asked to record wound infection and any therapeutic intervention. Patients were given bacteriology swabs and asked to return them to the laboratory if their wound was discharging. All patients reviewed at 6 weeks.</td>
</tr>
<tr>
<td></td>
<td></td>
<td><strong>Grading system:</strong> none</td>
<td>Results were presented separately according to the criteria used in the definition (e.g. surgeon’s diagnosis, purulent discharge, wound abscess).</td>
</tr>
</tbody>
</table>

RCT, randomised controlled trial

continued
### TABLE 23 contd  Prospective studies of surgical wound infection

<table>
<thead>
<tr>
<th>Study</th>
<th>Study design, sample size, aim</th>
<th>Definition, grading system, cited references</th>
<th>Data collection, postdischarge surveillance</th>
</tr>
</thead>
</table>
| Taylor et al., 1998,
Canada | Prospective postdischarge surveillance study of abdominal and vaginal hysterectomy (n = 763)  
**Aim:** to determine the frequency of SSI development after discharge from the hospital after abdominal or vaginal hysterectomy and the frequency of use of antimicrobial prophylaxis | **Definition:** 1992 CDC  
**Reference to definition:** Horan and co-workers  
**Grading system:** none | **Hospital:** not specified. The aim was to assess patients postdischarge  
**Community:** surgeons were asked to answer the following questions:  
1. Has the patient been seen in follow-up?  
2. Has SSI developed after discharge, as defined by the CDC criteria?  
3. Was the infection superficial, deep or involving organ space?  
4. Was there a need for re-admission to hospital?  
It was not possible to independently confirm the diagnosis of infection as records were held off-site. Post-discharge SSI outcome data were available for 99.6% of the sample. The authors state that there were very few single-institution published reports on the occurrence of SSIs after discharge in surgical gynaecology |
| Taylor et al., 1995,
Canada | Nested case–control study of multiple surgical procedures, including general, cardiac, vascular, neurological and orthopaedic (n = 4702)  
**Aim:** to determine the effect of SWI on the postoperative duration of hospital stay | **Definition:** 1992 CDC  
**Reference to definition:** Horan and co-workers  
**Grading system:** none | **Hospital:** direct inspection of incision  
**Community:** followed up until discharge or for 14 days |
| van Griethuysen et al., 1996,
Netherlands | Prospective before (n = 2905) and after (n = 2935) study in general and orthopaedic surgery  
**Aim:** to compare the results of ongoing surveillance of postoperative wound infection by an infection control nurse during a 9-month period before and after moving to a new theatre | **Definition:** 1988 CDC  
**Reference to definition:** Garner and co-workers  
**Grading system:** none | **Hospital:** infection data were collected by ward personnel, and interpreted and collated by an infection control nurse. Denominator data were obtained from the surgical register  
**Community:** follow-up at 1 month (general) and 1 year (orthopaedic implants) |
| Vegas et al., 1993,
Spain | Prospective cohort matched infected–uninfected study in general and digestive surgery (n = 1143)  
**Aim:** to estimate the excess hospital stay attributable to the nosocomial infections of patients from general and digestive surgery, and direct costs | **Definition:** 1988 CDC  
**Reference to definition:** Garner and co-workers  
**Grading system:** none | **Hospital:** a nurse trained in epidemiology recorded all the clinical data  
**Community:** not specified |

SSI, surgical site infection; SWI, surgical wound infection
<table>
<thead>
<tr>
<th>Study</th>
<th>Study design, sample size, aim</th>
<th>Definition, grading system, cited references</th>
<th>Data collection, postdischarge surveillance</th>
</tr>
</thead>
</table>
| Velasco et al., 1996,64 Brazil | Prospective cohort study of abdominal surgery for cancer (n = 236) | **Definition**: 1988 and 1992 CDC<sup>31,32</sup>  
**Reference to definition**: Garner and co-workers,<sup>31</sup> Horan and co-workers<sup>32</sup>  
**Grading system**: none | **Hospital**: infection control nurses conducted postoperative surveillance daily until discharge  
**Community**: not assessed. Survey of nosocomial infection rates |
| Velasco et al., 1998,63 Brazil | Prospective cohort study of multiple surgical procedures (n = 1205) | **Definition**: 1992 CDC<sup>32</sup>  
**Reference to definition**: Horan and co-workers<sup>32</sup>  
**Grading system**: none | **Hospital**: patients were visited 3–4 times weekly by trained infection control nurses  
**Community**: patients were not followed up after discharge. If a patient was re-admitted with infection, an association with the original operation was investigated |
| Vuorisalo et al., 1998,64 Finland | RCT of antibiotic prophylaxis in CABG surgery (n = 884) | **Definition**: 1988 and 1992 CDC<sup>31,32</sup>  
**Reference to definition**: Garner and co-workers,<sup>31</sup> Horan and co-workers<sup>32</sup>  
**Grading system**: none | **Hospital**: infection data were collected daily until discharge by a single cardiac surgeon. Later, the data were classified according to the CDC criteria by an infectious disease physician  
**Community**: all patients were given a questionnaire at discharge and requested to complete it during the 1-month follow-up visit. Postdischarge data could not be classified on CDC criteria. A copy of the questionnaire was printed in the article. Results were presented according to the 1992 CDC definitions |
| Vuorisalo et al., 1998,65 Finland  
(published in conjunction with the above article) | Risk analysis of prospective data on CABG surgery (n = 884) | **Definition**: 1992 CDC<sup>32</sup>  
**Reference to definition**: Horan and co-workers<sup>32</sup>  
**Grading system**: none | **Hospital**: infection data were collected daily until discharge by a single cardiac surgeon. Later, data were classified by an infectious disease physician  
**Community**: all patients were given a questionnaire at discharge and requested to complete it during the 1-month follow-up visit. Postdischarge data could not be classified on CDC criteria. However, the prescription of antibiotics for wound infection after discharge was used as a surrogate for postdischarge wound infection |

RCT, randomised controlled trial; SSI, surgical site infection  
continued
<table>
<thead>
<tr>
<th>Study</th>
<th>Study design, sample size, aim</th>
<th>Definition, grading system, cited references</th>
<th>Data collection, postdischarge surveillance</th>
</tr>
</thead>
</table>
| Weiss et al., 1999, USA | Retrospective analysis of prospectively collected SSI surveillance data (n = 20,007 procedures) | Definition: 1992 CDC Ref: Horan and co-workers<sup>32</sup>  
Reference to definition: Horan and co-workers<sup>32</sup>  
Grading system: none  

Aim: to examine the impact of antibiotic restriction policies on the microbiological aspects of SSIs for 6 years, to disseminate SSI rates and examine the effect of post-discharge surveillance | Hospital: all wound infections were identified through a process of review of medical records, morbidity and mortality records, culture reports and quarterly questionnaires completed by surgeons. All wound infections were followed up by a nurse clinician who acted as a consultant/data manager  
Community: follow-up was attempted at 30 days for all patients and up to 1 year for implant surgery: (1) clinic appointment follow-up, (2) telephone calls to patients lost to follow-up, (3) contact with regional ICPs regarding patients seen at other institutions |
| Wikblad and Anderson, 1995, Sweden | RCT of wound dressings in CABG or valve surgery (n = 250)  
Aim: to assess clinical aspects of a semi-occlusive hydroactive dressing and an occlusive hydrocolloid dressing in comparison to a conventional absorbent non-occlusive dressing | Definition: as per grading system below  
Reference to definition: none  
Grading system: assessed for redness (0–3 scale), degree of healing (0–3 scale)  
Redness: 0, no redness; 1, slight redness; 2, excessive redness  
Wound healing scale: 1, well healed (wound edges well together; a gap of less than 5% of the entire length of the incision allowed, with no or slight redness); 2, partially healed (gaps > 5% but < 20% of the whole length of the incision, with slight to excessive redness); 3, poorly healed (gaps > 20% of the entire length of the incision, with excessive redness) | Hospital: 216 (86%) patients were evaluated during their hospital stay, for 5 consecutive days. Five nurses were trained to evaluate wounds. On day 5, a culture and a colour photograph (standardised distance and angle) were taken and used to assess redness and degree of wound healing  
Community: At week 4 post-operatively, public health nurses assessed incisions and completed the protocol and mailed it to the clinic |
| Wong et al., 1997, Australia | Prospective study of leg wound infections following CABG (n = 152)  
Aim: to identify risk factors for infection and possible benefits of saline lavage in saphenous vein graft for coronary revascularisation | Definition: presence of purulent drainage and erythema of wound edges  
Reference to definition: none  
Grading system: none | Hospital: assessed at day 5  
Community: assessed at week 6. No other details given |
| Yalcin et al., 1995, Turkey | Prospective study of multiple surgical procedures (n = 4146)  
Aim: to determine the rate of infection in a Turkish hospital and factors that influence this rate | Definition: 1988 CDC  
Reference to definition: Garner and co-workers, Lim et al., Ayiffe and co-workers, Glenister and co-workers  
Grading system: none | Hospital: observed daily by an infection control nurse with a physician from the Infection Control Committee  
Community: wounds were followed up for at least 28 days |

<sup>RCT</sup>, randomised controlled trial; <sup>SSI</sup>, surgical site infection
## TABLE 24 Validation studies of surgical wound infection

<table>
<thead>
<tr>
<th>Study</th>
<th>Study design, sample size, aim</th>
<th>Standard and gold practice</th>
<th>Validity and reliability</th>
</tr>
</thead>
<tbody>
<tr>
<td>Bitkover et al., 1996, Sweden</td>
<td>Prospective study of patients undergoing cardiac surgery ( n = 29 ) patients: 24 with signs and symptoms suggesting postoperative infection; 5 control patients</td>
<td><strong>Standard</strong>: objective tests using (^{99m})Tc labelled monoclonal granulocyte antibody with SPECT</td>
<td><strong>Sensitivity</strong>: of 24 patients with signs and symptoms of infection, 7 had pathologic scans; 2 of which were due to superficial wound infection</td>
</tr>
<tr>
<td></td>
<td></td>
<td><strong>Assessors</strong>: scans performed and evaluated by an experienced nuclear medicine physician (not blinded)</td>
<td><strong>Specificity</strong>: of the 17 negative scans in the symptomatic group, 9 patients were diagnosed with infection (but none of these had wound infection). No false-positive results</td>
</tr>
<tr>
<td></td>
<td></td>
<td><strong>Assessor agreement</strong>: not given</td>
<td></td>
</tr>
<tr>
<td>Breidenbach and Trager, 1995, USA</td>
<td>Prospective study to compare ( n = 50 )</td>
<td><strong>Standard</strong>: quantitative cultures compared with (a) swab culture, (b) mechanism of injury, (c) severity of fracture, (d) wound position</td>
<td><strong>Accuracy</strong>: not all results were replicated as sensitivity, specificity, PPVs and NPVs were given for the five separate tests ( n = 25 ) results. The test with the highest utility and validity was quantitative culture</td>
</tr>
<tr>
<td></td>
<td></td>
<td><strong>Gold standard</strong>: quantitative culture ( (1 \text{ g of tissue}) ) assessed for degree of infection; reported in colonies per gram of tissue at 24 and 48 hours</td>
<td><strong>Quantitative culture</strong>: sensitivity 89%, specificity 95%, PPV 89%, NPV 95%</td>
</tr>
<tr>
<td></td>
<td></td>
<td><strong>Swab culture</strong>: sensitivity 83%, specificity 20%, PPV 38%</td>
<td></td>
</tr>
<tr>
<td>Byrne et al., 1988, UK</td>
<td>Validation study of ASEPSIS in 100 patients undergoing general and vascular surgery</td>
<td>Two independent observers (one nursing sister and one surgical registrar) assessed 100 wounds over a 4-week period</td>
<td>Scores ranged from 0 to 42 (nursing sister) and 0 to 41 (surgical registrar). The (mean difference –0.1, correlation coefficient 0.96, coefficient of repeatability 3.4)</td>
</tr>
</tbody>
</table>

NPV, negative predictive value; PPV, positive predictive value; SPECT, single photon emission computed tomography

continued
TABLE 24 contd Validation studies of surgical wound infection

<table>
<thead>
<tr>
<th>Study</th>
<th>Study design, sample size, aim</th>
<th>Standard and gold practice</th>
<th>Validity and reliability</th>
</tr>
</thead>
<tbody>
<tr>
<td>Cardo et al., 1993, USA</td>
<td>Prospective validation study over two periods (December 1990 to October 1991; May 1992 to April 1992) in a university-affiliated regional medical centre (n = 925 patients). All inpatient gastrointestinal procedures (except oesophageal, rectal, anal), amputations, breast, herniorrhaphies</td>
<td><strong>Standard</strong>: surveillance was performed by three ICPs who reviewed notes, charts, records and results and undertook staff consultation. Performed on days 4 and 7 postoperatively and weekly until discharge. ICPs examined a wound if a discussion with personnel did not determine a SWI. ICPs were given a 1 month pilot period prior to the validation study.</td>
<td><strong>Accuracy of classification by ICP surveillance</strong>: First period: sensitivity 84.8% (95% CI, 76 to 92), specificity 99.8% (95% CI, 99 to 100) Second period: sensitivity 92.3% (95% CI, 62 to 100), specificity 99% (95% CI, 94 to 100)</td>
</tr>
<tr>
<td></td>
<td><strong>Gold standard</strong>: hospital epidemiologist (first period) and assistant hospital epidemiologist (second period) examined wounds daily in addition to the above standard surveillance</td>
<td><strong>Gold standard</strong>: hospital epidemiologist (first period) and assistant hospital epidemiologist (second period) examined wounds daily in addition to the above standard surveillance</td>
<td><strong>Accuracy of surveillance related to experience. Sensitivity dropped when a new ICP started, and then recovered; it was 100% for the final 3 months of the study. The authors concluded that false reporting of wound infections would not be a problem, and that direct observation was not necessary for accurate identification of a SWI.</strong></td>
</tr>
</tbody>
</table>

Aim: to determine the sensitivity and specificity of standard infection control surveillance techniques for identifying surgical wound infections

Accuracy was low (62%) for the first 16 trauma cases evaluated by the CNs, and thus a second education session was held; accuracy increased to 91.2%

NA/CN for four classifications for 603 procedures: 56.7% (95% CI, 53 to 61)

NA/CN for two classifications (clean/clean–contaminated versus contaminated/dirty–infected) for 603 procedures: 75% (95% CI, 72 to 79)

Errors in classification were random and were not confined to one particular category. Classification of trauma surgery was more difficult, often due to multiple procedures through same incision

CI, confidence interval; CN, circulating nurse; NA, nurse anaesthetist; PO, physician observer; SWI, surgical wound infection

continued
### TABLE 24 contd Validation studies of surgical wound infection

<table>
<thead>
<tr>
<th>Study</th>
<th>Study design, sample size, aim</th>
<th>Standard and gold practice</th>
<th>Validity and reliability</th>
</tr>
</thead>
<tbody>
<tr>
<td>Ehrenkranz et al., 1995; USA</td>
<td>Retrospective review based on standardised medical record audits conducted every 1–2 years. 16 ICPs from 14 hospitals were evaluated, giving total of 28 ICP-years of observation. <strong>Study aim</strong>: to compare the accuracy of classification by ICPs of operative site infection in FCIC hospitals, in two time periods, and to estimate the effect of duration of surveillance experience on that accuracy.</td>
<td><strong>Standard</strong>: surveillance performed by ICPs using FCIC methods and criteria for infection. <strong>Gold standard</strong>: medical record reviewers examined the records of all patients classified by ICPs as infected to identify false-positive classifications. Reviewers also examined a sample of 100 records from patients classified as non-infected to identify false-negative classifications.</td>
<td><strong>Agreement</strong>: For each ICP: sensitivity estimates ≥ 80% and specificity estimates ≥ 97% were considered satisfactory. All ICPs: sensitivity 85–100%, specificity 97–100%. There was a wide variation in sensitivity, although this improved with experience.</td>
</tr>
<tr>
<td>Ehrenkranz et al., 1995, USA</td>
<td>A case–control study to evaluate false-positive diagnoses in laminectomy patients at one community hospital (n = 18; 18 matched controls). <strong>Study aim</strong>: to describe the outcome of an epidemiological investigation of apparently over-reported infections in the practice of one surgeon in a community hospital.</td>
<td><strong>Standard</strong>: comparison of patients as categorised by one physician. <strong>Gold standard</strong>: independent external investigation. Control patients with wound purulence or cellulitis were recorded as documented OSI, and control patients with a doctor’s diagnosis as the sole criterion were recorded as presumptive.</td>
<td><strong>Accuracy</strong>: Frequency of adverse events within operative site: documented OSI patients 83%, presumptive OSI patients and controls 16.7%. The similarity of the frequency of adverse events in presumptive OSI patients and controls suggests that there was excess diagnosis of SSI due to incorrect diagnosis.</td>
</tr>
<tr>
<td>Gur et al., 1998, Israel</td>
<td>Validation study of CT as the gold standard for imaging of postoperative sternal wound infection in plastic and reconstructive surgery patients at one medical centre (n = 203). <strong>Study aim</strong>: to evaluate the accuracy and role of CT in diagnosing the extent of infectious complications following sternotomy.</td>
<td><strong>Standard</strong>: CT. <strong>Gold standard</strong>: all available clinical and radiological data (multiple imaging methods). Correlation between CT and intraoperative clinical findings confirmed by histopathological studies.</td>
<td><strong>Accuracy</strong>: values were different for detecting different pathologies. Overall values for the use of CT in the detection of soft tissue and sternal mediastinitis: sensitivity 93.5%, specificity 81.7%. Values for diagnoses of anterior sternal plate infections: sensitivity 96.2%, specificity 92.8%. Values for diagnoses of posterior sternal plate infections: sensitivity 85.1%, specificity 92.8%.</td>
</tr>
<tr>
<td>Lima et al., 1993, Brazil</td>
<td>Prospective validation of an alternative nosocomial surveillance method in all patients with an identified risk factor for nosocomial infection as identified by a form completed by resident physicians (n = 376 cases). <strong>Study aim</strong>: to describe the performance of a new method of selective chart surveillance for nosocomial infections based on risk factors identified by physicians.</td>
<td><strong>Standard</strong>: selective surveillance by infection control nurses, by reviewing medical notes and other recorded information. Results were compared to the gold standard. <strong>Gold standard</strong>: two physician specialists reviewed the charts of all hospitalised patients on three occasions (after 1, 6 and 24 months of surveillance).</td>
<td><strong>Accuracy</strong>: Surveillance method: sensitivity 74% (95% CI, 54 to 93), specificity 99.7% (95% CI, 99.2 to 100), PPV 93% (95% CI, 81 to 100), NPV 99% (95% CI, 97.5 to 99.8). Overall accuracy of nosocomial surveillance 98% (95% CI, 97.3 to 99.7).</td>
</tr>
</tbody>
</table>

FCIC, Florida Consortium for Infection Control; NPV, negative predictive value; OSI, operative site infection; PPV, positive predictive value; SSI, surgical site infection

*continued*
TABLE 24 contd  Validation studies of surgical wound infection

<table>
<thead>
<tr>
<th>Study</th>
<th>Study design, sample size, aim</th>
<th>Standard and gold practice</th>
<th>Validity and reliability</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mitchell et al.,</td>
<td>Prospective study in 1360 patients undergoing major elective surgery</td>
<td><strong>Standard</strong>: patient assessment of the presence of a SWI</td>
<td>641 forms were returned</td>
</tr>
<tr>
<td>1999, Australia</td>
<td><strong>Aim</strong>: to evaluate postdischarge surgical wound surveillance using mailed-out questionnaires completed independently by patients and surgeons; and to compare the incidence and outcomes of wound infection that develops prior to patients’ discharge with those developing after hospital discharge</td>
<td><strong>Gold standard</strong>: physician assessment of the presence of a SWI, defined by purulent drainage</td>
<td>Infection was present in 51 and absent in 565 patients, as assessed by both patients and surgeons (κ = 0.73). Eight infections detected by surgeons were not detected by patients; 25 infections reported by patients were not diagnosed by surgeons</td>
</tr>
<tr>
<td>Oates and Payne, 1994,</td>
<td>Validation of diagnostic imaging test for postoperative infection in patients who had undergone sternotomy (n = 49)</td>
<td><strong>Standard</strong>: 50 postoperative scans performed in 49 patients. The findings were correlated with the sites of postoperative infection</td>
<td>Accuracy: correlation of $^{111}$In labelled white blood cell scans with infected sites</td>
</tr>
<tr>
<td>USA</td>
<td><strong>Aim</strong>: to examine the diagnostic role and scope of $^{111}$In labelled white blood cell scintigraphy in a selected yet diverse group of patients with complicated postoperative cardiothoracic conditions</td>
<td><strong>Gold standard</strong>: review of medical notes, infection records, laboratory results and results from a minimum clinical follow-up of 1 year</td>
<td>All sites: sensitivity 86%, specificity 97%, accuracy 95%</td>
</tr>
<tr>
<td>Poulson and Meyer, 1996,</td>
<td>External validation study of a hospital surveillance system over two study periods separated by 6 weeks (n = 1002 patients). Comprised 10% of surgical inpatients in Denmark (3 university surgical departments, 9 regional country hospital departments, 3 local hospitals)</td>
<td><strong>Standard</strong>: surgeon completed surgical details on a registration form immediately after the operation. Any infection detected was recorded by a doctor on a second form, stored in the records and subsequently entered in the routine electronic hospital surveillance database. The 1988 CDC definition of surgical wound infection was used</td>
<td>Accuracy: the total sensitivity of the routine hospital surveillance system was 26%, as compared with observation by researchers. The main problems were with completion of the basic registration and infection forms. The authors state that the Danish system is a low-cost model compared to US systems that rely on specialised infection control teams</td>
</tr>
<tr>
<td>Denmark</td>
<td><strong>Aim</strong>: (1) to perform an external validation of the hospital routine surveillance of SWI through two bedside prevalence surveys in 15 surgical and gynaecological departments; (2) postdischarge surveillance of a cohort of patients</td>
<td><strong>Gold standard</strong>: for the validation study, bedside inspection of all wounds was also performed by one of the authors</td>
<td></td>
</tr>
</tbody>
</table>

SWI, surgical wound infection

continued
## TABLE 24 contd Validation studies of surgical wound infection

<table>
<thead>
<tr>
<th>Study</th>
<th>Study design, sample size, aim</th>
<th>Standard and gold practice</th>
<th>Validity and reliability</th>
</tr>
</thead>
<tbody>
<tr>
<td>Sands et al., 1996,197 USA</td>
<td>Prospective comparison of postdischarge surveillance methods on patients undergoing non-obstetric procedures at one major hospital (n = 5572 procedures)</td>
<td><strong>Standard</strong>: patients received a single-page questionnaire that contained yes/no tick boxes for questions relating to signs and symptoms of infection on days 25–32 postoperatively. Surgeons received a form every 4 weeks that listed operated cases from the previous 408 weeks. They were asked to specify whether an SSI was possible, definite, not present or don’t know. <strong>Gold standard</strong>: the performance characteristics of patients and surgeons were compared with record review procedures performed by two infectious disease physicians who independently reviewed all documented records.</td>
<td><strong>Accuracy</strong>: Patient response rate 1799/5572 (33.4%) Positive patient response: sensitivity 28%, PPV 36%; sensitivity when unreturned questionnaires excluded, 68% Surgeon response rate: 4420/5572 (79%) If possible, surgeon interpreted SSI as negative: sensitivity 15%, PPV 28% If possible, surgeon interpreted SSI as positive: sensitivity 24%, PPV 19%</td>
</tr>
<tr>
<td>Sands et al., 1999,197 USA</td>
<td>Prospective comparison of patients undergoing non-obstetric procedures at one major hospital (n = 4086 procedures)</td>
<td><strong>Standard</strong>: design and use of an algorithm on statistical software using computerised data from claims, pharmacy records and automated medical records to identify SSIs. <strong>Gold standard</strong>: compared against SSI data collected in previous study197.</td>
<td>Overall for the model: sensitivity 74%, specificity 98%, PPV 48% Other models with higher sensitivity (higher cost for false-negative results) had corresponding decrements in specificity and PPV. Models created by restricting types of data sources did not perform as well compared to when all data sources were available, and it was not possible to create a model with a sensitivity &gt; 80% or a PPV &gt; 35% without information on outpatient tests and diagnoses. Thus automated postdischarge medical records detected SSIs with better sensitivity and specificity than did patient or surgeon surveys.</td>
</tr>
<tr>
<td>Seaman and Lammers, 1991,197 USA</td>
<td>Validation of patients’ ability to self-diagnose wound infections in patients with lacerations sutured at an emergency department</td>
<td><strong>Standard</strong>: patient assessment based on response to a standardised interview. <strong>Gold standard</strong>: medical assessment and completion of a physician assessment form by assessors. Assessors were nurse practitioners (n = 3), physicians (n = 51) and a physician’s assistant (n = 1).</td>
<td>Values given below are for patient diagnosis compared with the gold standard. Infection: sensitivity 0.52, specificity 0.92, PPV 0.26, NPV 0.97, accuracy 0.91. Purulence: sensitivity 0.47, specificity 0.97, PPV 0.39, NPV 0.98, accuracy 0.96. Redness: sensitivity 0.68, specificity 0.80, PPV 0.21, NPV 0.97, accuracy 0.79. Swelling: sensitivity 0.22, specificity 0.92, PPV 0.13, NPV 0.95, accuracy 0.88. Warmth: sensitivity 0.58, specificity 0.81, PPV 0.16, NPV 0.97, accuracy 0.80. Tenderness: sensitivity 0.42, specificity 0.84, PPV 0.35, NPV 0.88, accuracy 0.77.</td>
</tr>
</tbody>
</table>
TABLE 24 contd  Validation studies of surgical wound infection

<table>
<thead>
<tr>
<th>Study</th>
<th>Study design, sample size, aim</th>
<th>Standard and gold practice</th>
<th>Validity and reliability</th>
</tr>
</thead>
<tbody>
<tr>
<td>Severijnen et al., 1997, Netherlands</td>
<td>Validation study of (a) patient-finding by comparing ICP admission numbers with those from the national register; and (b) case-finding by local ICPs. Eight hospitals in central Netherlands covering general gynaecology and orthopaedic surgery</td>
<td><strong>Standard</strong>: patient-finding done by registration of all admissions by ICPs and compared with a national data source to examine the completeness of the two registrations. Case-finding (SWI) of patients done by local ICPs. Data were collected by ICPs who visited wards twice weekly, checking charts, records and reports for signs of infection. Patients with nosocomial infection were registered and further data obtained</td>
<td>Patient-finding; ICPs registered 1946 admissions (91% present on LMR) Case-finding by local ICP surveillance (n = 316 patients): sensitivity 87.5%, specificity 98.6%</td>
</tr>
<tr>
<td>Smyth et al., 1997, Northern Ireland</td>
<td>Comparison of accuracy of manual versus automated data entry using 100 SWI surveillance questionnaires</td>
<td><strong>Standard</strong>: manual versus automated data entry using Formic optical scanning technology</td>
<td>Automated data entry: accuracy 99.98% (&lt; 0.2 errors per 1000 responses) Manual data entry: accuracy 98.76% (12.4 errors per 1000 responses)</td>
</tr>
<tr>
<td>Wilson et al., 1986, UK</td>
<td>Validation study of the ASEPSIS wound scoring system in cardiac patients</td>
<td>Two independent observers assessed 51 sternal wounds and 34 leg wounds from 51 patients</td>
<td>The scores for sternal wounds ranged from 0 to 22 points and from 0 to 18 points from each observer, and those for leg wounds from 0 to 13 and 0 to 15 points, respectively. The mean difference was 0 and 0.1 points for sternal and leg wounds, respectively. The coefficient of repeatability was 4.1 points for sternal wounds and 3.2 for leg wounds</td>
</tr>
</tbody>
</table>

SWI, surgical wound infection
<table>
<thead>
<tr>
<th>Study</th>
<th>Study design, sample size, aim</th>
<th>Standard and gold practice</th>
<th>Validity and reliability</th>
</tr>
</thead>
<tbody>
<tr>
<td>Wilson et al., 1990,119 UK</td>
<td>Prospective validation study in 1029 surgical patients with wounds &gt; 2 cm</td>
<td><strong>Standard</strong>: ASEPSIS scores above 10, 20, 30 and 40 points were compared with the definitions given below. <strong>Gold standard</strong>: Cruse and Foord119 definition; Leigh187 grading scale. Other variables related to a change in management were also used, including length of hospital stay and antibiotic usage.</td>
<td>Daily scores were allocated during the first week and further points were given for other criteria satisfied in the first 2 months after surgery. Wounds were assessed daily by nursing staff, including at weekends, and scores were checked every 1–3 days by one of two authors. A total of 48 sensitivity and specificity values were given in the published article. These are not reproduced here, but are discussed in the text in chapter 5 of this review.</td>
</tr>
<tr>
<td>Wilson et al., 1998,127 UK</td>
<td>Prospective study over 2 months in 1993 and 1995; wounds were examined by the same observer</td>
<td><strong>Standard</strong>: ASEPSIS and SWAS <strong>Gold standard</strong>: the 1988 CDC31 and NPS10 definitions</td>
<td>The CDC and NPS definitions did not differ significantly from each other. The ASEPSIS and SWAS systems did not differ significantly from each other. 44–47% of wounds infected as per gold standard definitions were classed as a disturbance of healing by ASEPSIS. All methods of assessment were labour intensive.</td>
</tr>
<tr>
<td>Wischenewski et al., 1998,143 Gastmeier, 1998,144 Germany</td>
<td>Validation study of the investigators at the start and end of the first German prevalence study of nosocomial infection in patients undergoing trauma, abdominal and gynaecology/obstetric surgery (72 randomly selected German hospitals, 14,966 patients)</td>
<td><strong>Standard</strong>: four physician investigators recorded data for the survey and underwent a validation period in one hospital. Two types of validation (bedside assessment and case studies) were performed in two phases, at the beginning and the end of the prevalence survey, where each investigator independently examined the same 100 patients</td>
<td>Overall for investigators: sensitivity 89%, specificity 99.3%. Bedside validation (n = 200 patients): sensitivity 89.0%, specificity 99.3%. Validation by case studies (n = 60): sensitivity 95.6%, specificity 92.8%.</td>
</tr>
<tr>
<td>Yokoe et al., 1998,173 USA and Israel</td>
<td>Prospective comparison of routine surveillance compared with antibiotic exposure in CABG patients (n = 3887)</td>
<td><strong>Standard</strong>: SSIs identified by antibiotic exposure <strong>Gold standard</strong>: SSIs identified by conventional prospective surveillance from two national systems: (1) USA – routine SSI surveillance by ICPs, by chart review, laboratory results and ward rounds; (2) Israel – based on the Israeli Study of Surgical Infections whereby nurse epidemiologists perform surveillance and postdischarge follow-up</td>
<td>USA: a minimum antibiotic interval of 9 days following a lag period of 1 day gave the best (extrapolated) combination of sensitivity and specificity (sensitivity 95%, specificity 85%). The PPV of antibiotic exposure for detecting SSIs was 28% and that of antibiotic exposure for detecting nosocomial infections was 60% in Israel with an antibiotic exposure threshold value of at least 9 days of postoperative antibiotics the sensitivity of detecting SSIs was 87%, the specificity 82% and the PPV 31%.</td>
</tr>
</tbody>
</table>

*PPV, positive predictive value; SSI, surgical site infection; SWAS, Southampton Wound Assessment Scale*
### TABLE 25  Studies of anastomotic leak

<table>
<thead>
<tr>
<th>Study</th>
<th>Type of surgery, aim of study</th>
<th>Definition, clinical features, investigations</th>
</tr>
</thead>
<tbody>
<tr>
<td>Abo et al., 1996, Japan</td>
<td>Surgery: thoracic oesophagectomy (n = 564)</td>
<td><strong>Definition:</strong> none</td>
</tr>
<tr>
<td></td>
<td><strong>Aim:</strong> to evaluate the significance of improvements in surgical technique by studying the long-term outcome of 500 patients</td>
<td><strong>Clinical features:</strong> not described</td>
</tr>
<tr>
<td></td>
<td><strong>Investigations:</strong> daily postoperative tissue oxygen tension (P&lt;sub&gt;O&lt;/sub&gt;&lt;sub&gt;2&lt;/sub&gt;) levels using a sensor placed at the anastomotic site. Blood flow was also measured in the upper, middle and lower regions of the gastric tube using laser Doppler velocimetry to measure the relationship between blood flow and leak</td>
<td></td>
</tr>
<tr>
<td>Ah Chong et al., 1996, Hong Kong</td>
<td>Surgery: multiple procedures (oesophagectomy; oesophageal bypass; gastrectomy; biliary, hepatic, pancreatic resection; colorectal surgery) (n = 180)</td>
<td><strong>Definition:</strong> none</td>
</tr>
<tr>
<td></td>
<td><strong>Aim:</strong> a prospective audit of single-layer continuous anastomosis in gastrointestinal surgery</td>
<td><strong>Clinical features:</strong> not described, although the authors state that investigations were performed if clinically indicated</td>
</tr>
<tr>
<td></td>
<td><strong>Investigations:</strong> surgery type divided into: group I, upper gastrointestinal; group II, hepatobiliary; group III, colorectal. Routine retrograde cholangiogram was performed for biliary/choledochoenteric anastomosis. No routine postoperative contrast studies for oesophageal, gastric or colonic anastomoses, unless clinically indicated. Gastrografin swallow for suspected upper gastrointestinal leaks</td>
<td></td>
</tr>
<tr>
<td>Ambrossetti et al., 1994, Switzerland</td>
<td>Surgery: colorectal surgery (n = 200)</td>
<td><strong>Definition:</strong> none</td>
</tr>
<tr>
<td></td>
<td><strong>Aim:</strong> a prospective study to report on the experience with the first 200 anastomoses performed on an elective basis, using a standardised technique and systematic use of Doppler ultrasound to assess good vascularisation of the intestinal edges to be Anastomosed</td>
<td><strong>Clinical features:</strong> clinical leaks reported, but no details given of features or assessment</td>
</tr>
<tr>
<td></td>
<td><strong>Investigations:</strong> routine imaging with Gastrografin enemas between the elective days 9 and 11 postoperatively before oral intake was commenced</td>
<td></td>
</tr>
<tr>
<td>Anikin et al., 1997, Northern Ireland</td>
<td>Surgery: oesophagectomy (n = 113)</td>
<td><strong>Definition:</strong> none</td>
</tr>
<tr>
<td></td>
<td><strong>Aim:</strong> to describe a technique of oesophageal resection that allows reliable and safe access to the chest, abdomen and neck</td>
<td><strong>Clinical features:</strong> none</td>
</tr>
<tr>
<td></td>
<td><strong>Investigations:</strong> routine radiological investigation of the anastomosis with water-soluble contrast on the day 6 postoperatively before oral intake was commenced</td>
<td></td>
</tr>
<tr>
<td>Biondo et al., 1997, Spain</td>
<td>Surgery: colonic resection for peritonitis or obstruction (n = 212)</td>
<td><strong>Definition:</strong> none</td>
</tr>
<tr>
<td></td>
<td><strong>Aim:</strong> to present the experience of left-sided large bowel conditions requiring emergency surgery, reporting on a series of 63 patients who had primary anastomosis and comparing the results in patients with peritonitis with those patients with obstruction</td>
<td><strong>Clinical features:</strong> clinical leaks reported, but no details given of features or assessment</td>
</tr>
<tr>
<td></td>
<td><strong>Investigations:</strong> one small bowel fistula was confirmed by contrast radiography</td>
<td></td>
</tr>
<tr>
<td>Bokey et al., 1995, Australia</td>
<td>Surgery: resection for colorectal cancer (n = 1846)</td>
<td><strong>Definition:</strong> anastomotic leaks were classified as subclinical and clinical</td>
</tr>
<tr>
<td></td>
<td><strong>Aim:</strong> to prospectively document and review the results to form a basis for evaluating new techniques (endosurgery and laparoscopic bowel resection)</td>
<td><strong>Clinical features:</strong> clinical and radiological confirmation. A significant clinical (general) leak was one that necessitated abdominal re-operation. A subclinical (local) leak was defined as one demonstrated by limited Gastrografin enema or one that resulted in an abscess that discharged either spontaneously or following minor surgical drainage</td>
</tr>
<tr>
<td></td>
<td><strong>Investigations:</strong> Gastrografin enema, but it is unclear whether this was conducted routinely</td>
<td></td>
</tr>
</tbody>
</table>

continued
Appendix 3

### TABLE 25 contd Studies of anastomotic leak

<table>
<thead>
<tr>
<th>Study</th>
<th>Type of surgery, aim of study</th>
<th>Definition, clinical features, investigations</th>
</tr>
</thead>
<tbody>
<tr>
<td>Bottger et al., 1999,245</td>
<td>Surgery: pancreaticoduodenectomy (n = 221)</td>
<td><strong>Definition:</strong> there was an attempt to distinguish between a pancreatic fistula and insufficient pancreatic anastomosis. Pancreatic fistula was identified in the presence of an amylase concentration in the drainage fluid of &gt; 2000 units/l.</td>
</tr>
<tr>
<td></td>
<td><strong>Aim:</strong> a prospective case–control to evaluate risk factors for morbidity and mortality after pancreaticoduodenectomy</td>
<td><strong>Clinical features:</strong> not described</td>
</tr>
<tr>
<td></td>
<td><strong>Investigations:</strong> see above</td>
<td></td>
</tr>
<tr>
<td>Bouillot et al., 1998,289</td>
<td>Surgery: minimally invasive laparoscopic colectomy (n = 50)</td>
<td><strong>Definition:</strong> none</td>
</tr>
<tr>
<td></td>
<td><strong>Aim:</strong> to evaluate prospectively the results in the first 50 patients in whom an elective laparoscopically assisted colonic resection was performed for sigmoid diverticulitis</td>
<td><strong>Clinical features:</strong> not described</td>
</tr>
<tr>
<td></td>
<td><strong>Investigations:</strong> one leak was diagnosed by radiographic study</td>
<td></td>
</tr>
<tr>
<td>Breen et al., 1998,261</td>
<td>Surgery: ileal pouch–anal anastomosis (n = 628)</td>
<td><strong>Definition:</strong> anastomotic separation was defined as a separation of the pouch–anal anastomosis, detected either clinically or by the retrograde water-soluble contrast study before closure of the ileostomy</td>
</tr>
<tr>
<td></td>
<td><strong>Aim:</strong> to determine the incidence of perineal complications after ileal pouch–anal anastomosis and the risk of pouch failure associated with these complications based on a review of a prospective registry</td>
<td><strong>Clinical features:</strong> not described</td>
</tr>
<tr>
<td></td>
<td><strong>Investigations:</strong> retrograde water-soluble contrast study before closure of the ileostomy</td>
<td></td>
</tr>
<tr>
<td>Burke et al., 1994,262</td>
<td>Surgery: colorectal surgery (n = 186)</td>
<td><strong>Definition:</strong> see below</td>
</tr>
<tr>
<td></td>
<td><strong>Aim:</strong> a prospective RCT to determine whether mechanical bowel preparation influences the incidence of anastomotic dehiscence following colorectal surgery</td>
<td><strong>Clinical features:</strong> anastomotic dehiscence was diagnosed clinically and suspected if there was: deterioration in the patient’s general condition; abdominal distension; diarrhoea or blood clot passed per anum; or signs of peritonitis. If necessary, leak was confirmed radiologically by using a water-soluble enema</td>
</tr>
<tr>
<td></td>
<td><strong>Investigations:</strong> in the first half of the series radiological leaks were routinely checked for in all patients with a colorectal anastomosis by administering a water-soluble contrast enema on day 7 postoperatively. Two of six leaks occurred on day 7 immediately after administration of a routine water-soluble contrast enema. These two complications led to a review of policy, and this investigation is now requested only when anastomotic leak is suspected on clinical grounds. Thus there are differences between the first and second half of this study</td>
<td></td>
</tr>
<tr>
<td>Choi et al., 1998,201</td>
<td>Surgery: oesophageal anastomosis (n = 40)</td>
<td><strong>Definition:</strong> none</td>
</tr>
<tr>
<td></td>
<td><strong>Aim:</strong> to evaluate the use of a neck drain after oesophagectomy in a prospective RCT</td>
<td><strong>Clinical features:</strong> neck wounds were inspected daily for evidence of haematoma or seroma formation. The drain was removed when output was &lt; 10 ml/day</td>
</tr>
<tr>
<td></td>
<td><strong>Investigations:</strong> on day 7 postoperatively, all patients had a water-soluble contrast study and endoscopy to detect subclinical as well as clinical leaks</td>
<td></td>
</tr>
<tr>
<td>Chou et al., 1996,246</td>
<td>Surgery: pancreaticoduodenectomy (n = 93)</td>
<td><strong>Definition:</strong> a pancreatic leak or fistula was defined as persistent drainage of ≥ 50 ml of amylase-rich fluid a day for &gt; 2 weeks. Gastric stasis was defined when the patient was unable to take liquids by mouth but had no evidence of abscess or leak for 10 days or postoperatively</td>
</tr>
<tr>
<td></td>
<td><strong>Aim:</strong> to elucidate the factors that influenced mortality among 93 patients who underwent pancreaticoduodenectomy for periampullary cancer, whether end-to-side pancreatic duct to jejun al mucosa anastomosis is better than end-to-end pancreaticojejunostomy in terms of morbidity and morality, and whether age &gt; 70 years is a contraindication for pancreaticoduodenectomy</td>
<td><strong>Clinical features:</strong> as above</td>
</tr>
<tr>
<td></td>
<td><strong>Investigations:</strong> as above</td>
<td></td>
</tr>
</tbody>
</table>

*RCT, randomised controlled trial*
### TABLE 25 contd Studies of anastomotic leak

<table>
<thead>
<tr>
<th>Study</th>
<th>Type of surgery, aim of study</th>
<th>Definition, clinical features, investigations</th>
</tr>
</thead>
</table>
| Cornwell et al., 1998, USA 263 USA | Surgery: penetrating abdominal/colonic trauma (n = 56) | **Definition:** the presence or absence of suture line disruption (where a colonic suture line existed)  
**Clinical features:** clinical and radiological confirmation. No description was given of how clinical features were assessed  
**Investigations:** all patients developing septic abdominal complications were evaluated for colonic suture line disruption either by surgical re-exploration, CT scan with intraluminal contrast, or Gastrografin enema |
| Craig et al., 1996, 214 UK | Surgery: oesophagogastrectomy (n = 100) | **Definition:** none  
**Clinical features:** not described  
**Investigations:** all patients underwent a barium swallow on the day 5 postoperatively to check for anastomosis. Patients with leak were treated (antibiotics, total parenteral nutrition) and the contrast examination repeated within a week of the first barium examination |
| Curry et al., 1998, 218 USA | Surgery: resectional gastric bypass in morbid obesity (n = 26) | **Definition:** none  
**Clinical features:** based on results which state that one patient developed signs of sepsis on day 2 postoperatively and was found to have a leak. The authors also refer to a later leak, where the patient presented on day 30 postoperatively with left pleural effusion, upper abdominal pain and signs of sepsis (found to be infection rather than leak)  
**Investigations:** CT conducted on the patient who presented on day 30. CT not done routinely |
| Davidson et al., 1999, UK 254 | Surgery: liver transplantation (n = 100) | **Definition:** biliary complications divided into: early (≤ 30 days after transplantation), intermediate (between day 30 postoperatively and 3 months) and late (> 3 months)  
**Clinical features:** not described  
**Investigations:** all patients routinely had endoscopic retrograde cholangiography 2 weeks (days 10–14) postoperatively. Results were reported by two experienced endoscopists as normal, leak or stricture. A patient was considered to have a biliary complication if cholangiography showed a stricture or leak, regardless of extent or severity |
| Dayton and Larsen, 1997, USA 290 | Surgery: ileal pouch–anal anastomosis (n = 510) | **Definition:** as below  
**Clinical features:** patients who presented with leak had the following: lower abdominal pain, tachycardia, fever and abdominal distension. Some presented with peritonitis, pneumoperitoneum and sepsis. Others presented with a walled-off abscess and a clinical picture of chronic, worsening lower abdominal pain  
**Investigations:** not done routinely, and not described |
| De Wever et al., 1996, 291 Belgium | Surgery: supravaleator pelvic exenteration | **Definition:** none  
**Clinical features:** the authors state that leak caused a pelvic abscess and sepsis in four patients  
**Investigations:** after 3–4 months the rectal anastomosis was considered healed when the patient was well clinically, endoscopically and radiologically  
continued |
TABLE 25 contd Studies of anastomotic leak

<table>
<thead>
<tr>
<th>Study</th>
<th>Type of surgery, aim of study</th>
<th>Definition, clinical features, investigations</th>
</tr>
</thead>
</table>
| Debus et al., 1999, Germany | Surgery: anterior resection of rectum (n = 75) | Definition: none  
Aim: a prospective study to evaluate perioperative mortality, anastomotic complications (such as bleeding and leak rate) in the biofragmentable anastomosis ring of the rectum. In addition, a long-term follow-up was undertaken to detect the rate of late anastomotic stenosis.  
Investigations: some barium studies were carried out on patients with complications (such as bleeding and prolonged bowel atony, which showed oedema of the anastomotic region). |
| Deen and Smart, 1995, UK | Surgery: colon resection (n = 53) | Definition: as below  
Aim: to evaluate prospectively the relative merits of continuous and interrupted single-layer colonic anastomoses.  
Clinical features: local or generalised peritonism, tachycardia, fever, frank faecal fistula, and anastomotic stricture.  
Investigations: anastomoses were not routinely examined radiographically to confirm their integrity. |
| Dehni et al., 1998, France | Surgery: colorectal resection (n = 258) | Definition: as below  
Aim: to report the experience of anastomotic complications in a series of 258 patients with cancers of 6–11 cm treated with colonic pouch–anal anastomosis or low colorectal anastomosis.  
Clinical features: clinical leak was defined as evidence of generalised or localised anastomotic complications.  
Investigations: all patients with a colonic pouch–anal anastomosis or low colorectal anastomosis and a defunctioning stoma underwent water-soluble contrast enema at 8–10 weeks before stoma closure. In addition, enema was performed in any patient with features suggesting a leak. |
| Deshmane and Shinde, 1994, India | Surgery: oesophagectomy (transhiatal; transthoracic) (n = 72) | Definition: anastomotic leak was defined as an asymptomatic small leak detected only by radiological study or a larger leak with a perianastomotic collection, which manifested clinically.  
Aim: to define the role of nutritional and technical factors in the pathogenesis of cervical oesophageo-gastric leak.  
Clinical features: not described  
Investigations: the anastomosis was routinely checked on day 10 postoperatively using a thin barium swallow. |
| Docherty et al., 1995, UK, The West of Scotland and Highland Anastomosis Study Group | Surgery: colorectal procedures (n = 652) | Definition: as below  
Aim: to compare surgical stapling with manual suturing in a multicentre, prospective, randomised study in a large population of patients undergoing both elective and emergency colorectal surgery.  
Clinical features: clinical leak was defined as a dehiscence at the anastomosis confirmed by re-operation or autopsy, the appearance of faecal material from drains, the development of a colocutaneous fistula, or the development of any systemic septic-associated local peritoneal signs in the postoperative period.  
Investigations: radiological assessment of anastomotic integrity was done using a water-soluble contrast enema performed between days 4 and 14 postoperatively. Any extravasation of the contrast medium detected on radiography was considered a radiological leak. |
| Evans et al., 1997, UK | Surgery: pancreatic surgery (n = 63) | Definition: as below  
Aim: a largely prospective study to analyse the medium-term outcome of surgery for chronic pancreatitis, with particular emphasis on the quality of life for patients based on the mortality and morbidity rates associated with surgery, relief of symptoms, analgesic use, employment, and long-term sequelae of pancreatic surgery.  
Clinical features: not described  
Investigations: the presence of a pancreatic fistula was defined by the production of > 50 ml/day abdominal fluid with an amylase content of > 1000 unit/l (normal value < 300 unit/l). |
### TABLE 25 contd Studies of anastomotic leak

<table>
<thead>
<tr>
<th>Study</th>
<th>Type of surgery, aim of study</th>
<th>Definition, clinical features, investigations</th>
</tr>
</thead>
<tbody>
<tr>
<td>Fernandez-Fernandez et al., 1995,203</td>
<td>Surgery: total gastrectomy (n = 101)</td>
<td>Definition: as per the classification of Csendes and co-workers(^{200}) (see below)</td>
</tr>
<tr>
<td></td>
<td><strong>Aim:</strong> a prospective observational analysis of the relationship between morbidity and the diameter of the cartridge used in oesophagojejunal anastomosis following total gastrectomy</td>
<td><strong>Clinical features:</strong> not described</td>
</tr>
<tr>
<td></td>
<td><strong>Investigations:</strong> routine imaging regardless of clinical signs. On day 7 postoperatively the sutures were checked by contrast radiography. All patients were examined endoscopically for the first time 6 months after surgery. Two types of fistula were identified according to Csendes' classification: type I, a local fistula with no dissemination through a fistulous track to the pleural or abdominal cavity, or the appearance of contrast material in any abdominal drain; type II, a leak with great dissemination to the pleural or abdominal cavity, with the appearance of contrast medium in any of the abdominal drains</td>
<td></td>
</tr>
<tr>
<td>Fernandez et al., 1996,204</td>
<td>Surgery: mechanical oesophageal anastomosis with fibrin glue (n = 86)</td>
<td>Definition: as per the classification of Csendes and co-workers(^{200}) (see below)</td>
</tr>
<tr>
<td></td>
<td><strong>Aim:</strong> to assess the effect of reinforcing fibrin glue on the anastomotic leak rate after mechanical oesophageal anastomosis</td>
<td><strong>Clinical features:</strong> not described</td>
</tr>
<tr>
<td></td>
<td><strong>Investigations:</strong> routine imaging regardless of clinical signs. On day 7 postoperatively the anastomosis was checked by contrast radiography. Two types of fistula were identified according to Csendes' classification: type I, a local fistula with no dissemination through a fistulous track to the pleural or abdominal cavity, or the appearance of contrast material in any abdominal drain; type II, a leak with great dissemination to the pleural or abdominal cavity, with the appearance of contrast medium in any of the abdominal drains</td>
<td></td>
</tr>
<tr>
<td>Fingerhut et al., 1994,267</td>
<td>Surgery: colon resection (n = 113)</td>
<td><strong>Definition:</strong> all images, but a perfectly regular anastomosis and uniform calibre of lumen were considered an anastomotic leak</td>
</tr>
<tr>
<td></td>
<td><strong>Aim:</strong> a prospective multicentre study to determine whether stapled infra-peritoneal colorectal anastomosis was associated with fewer postoperative (early and late) complications than were hand-sewn anastomoses</td>
<td><strong>Clinical features:</strong> overt leak was diagnosed by faecal matter in the drainage discharge, purulent discharge per anum, sinograms, re-operation or post-mortem examination. Covert leak was searched for routinely by means of a sodium benzoate enema on day 7 postoperatively</td>
</tr>
<tr>
<td></td>
<td><strong>Investigations:</strong> sodium benzoate enema performed on day 7 postoperatively</td>
<td></td>
</tr>
<tr>
<td>Fingerhut et al., 1995,268</td>
<td>Surgery: colon resection (n = 159)</td>
<td><strong>Definition:</strong> all images, but a perfectly regular anastomosis and uniform calibre of lumen were considered an anastomotic leak</td>
</tr>
<tr>
<td></td>
<td><strong>Aim:</strong> a prospective multicentre study to determine whether stapled anastomoses were associated with less morbidity and possibly less mortality in a standard setting of suprapерitoneal colorectal anastomosis for which both techniques, hand-sewn or stapled, were simple to perform</td>
<td><strong>Clinical features:</strong> overt leak was diagnosed by faecal matter in the drainage discharge, purulent discharge per anum, sinograms, re-operation or post-mortem examination. Covert leak was searched for routinely by means of a sodium benzoate enema on day 7 postoperatively</td>
</tr>
<tr>
<td></td>
<td><strong>Investigations:</strong> sodium benzoate enema performed on day 7 postoperatively</td>
<td></td>
</tr>
<tr>
<td>Flohr et al., 1996,273</td>
<td>Surgery: radical cystoprostatectomy and orthotopic bladder substitution via the ileal neobladder</td>
<td><strong>Definition:</strong> none</td>
</tr>
<tr>
<td></td>
<td><strong>Aim:</strong> to report on the morbidity and complications of a series of 306 men who received an ileal neobladder and in whom complete follow-up was available for an average of 4.2 years</td>
<td><strong>Clinical features:</strong> not described, but the authors do refer to insufficiency of the ileal anastomosis</td>
</tr>
<tr>
<td></td>
<td><strong>Investigations:</strong> a variety of examinations were conducted over the follow-up period (3-month intervals in years 1 and 2, 6-month intervals in years 3 and 4, annually thereafter). An excretory urogram (intravenous pyelogram) and voiding cystourethrogram were done after 1 year on all patients and when indicated</td>
<td></td>
</tr>
</tbody>
</table>
TABLE 25 contd  Studies of anastomotic leak

<table>
<thead>
<tr>
<th>Study</th>
<th>Type of surgery, aim of study</th>
<th>Definition, clinical features, investigations</th>
</tr>
</thead>
</table>
| Goel et al., 1995,
India | Surgery: oesophagectomy (n = 25)  
Aim: a prospective study comparing Gastrografin study with test feeding (done in the surgical ward using drinking water) in a series of 25 consecutive patients with cervical oesophagogastric anastomosis. The issues to be answered were: (1) are any anastomotic leaks missed by using 'test feeding' instead of Gastrografin study; and (2) if leaks are missed, do they lead to any complications? | Definition: none  
Clinical features: leak at neck site  
Investigations: routine imaging, regardless of clinical signs. All patients underwent Gastrografin study on day 5 postoperatively, followed by 'test feeding' with drinking water (more description is given on p. 68) |
| Gupta, 1996,
India | Surgery: transhiatal oesophagectomy (n = 250)  
Aim: to analyse the results for the first 250 unselected patients who underwent elective oesophagectomy for oesophageal carcinoma | Definition: none  
Clinical features: not described  
Investigations: contrast radiography on day 5. If anastomotic leak was detected, oral feeds were withheld. Contrast radiography was repeated weekly to see if the anastomotic leak had healed |
| Gupta et al., 1997,
India | Surgery: transhiatal oesophagectomy (n = 29)  
Aim: to describe the experience of transhiatal oesophagectomy without thoracotomy for benign disease | Definition: none  
Clinical features: not described  
Investigations: the integrity of anastomosis was checked by contrast radiography. If leak was detected, oral feeds were withheld and feeding through a Ryle's tube commenced. Contrast radiography was repeated weekly to confirm healing of the anastomotic leak |
| Gupta et al., 1998,
USA | Surgery: hepatobiliary surgery (n = 13)  
Aim: to compare the management and outcome of isolated bile duct injuries with bile duct and hepatic artery injuries | Definition: anastomotic leaks were identified as bile drainage from drains placed at the area of the anastomosis  
Clinical features: bile drainage, as above  
Investigations: anastomotic patency was evaluated primarily by biliary nucleotide scintigraphy and, where indicated, by transhepatic cholangiography |
| Hallbook et al.,
1996, Sweden and
USA | Surgery: low anterior colon resection (n = 100)  
Aim: to compare reconstruction with the traditional straight anastomosis and the colonic J-pouch anastomosis in a multicentre randomised study | Definition: symptomatic anastomotic leak was evident if any of the following was observed: evidence of abscess on a CT scan or ultrasound; discharge of pus either per anum or through a fistula; and the necessity of laparotomy or a transanal drainage procedure. The criteria of anastomotic stricture were fulfilled when a dilatation under anaesthesia was required  
Clinical features: as above  
Investigations: anastomotic integrity was confirmed by digital and endoscopic examination and also by contrast enema before closure of the temporary stoma in applicable patients |
| Hamanaka and Suzuki,
1994, Japan | Surgery: pancreatoduodenectomy (n = 48)  
Aim: to describe a simple leak-proof technique for performing end-to-side pancreatojejunostomy after pancreato-duodenectomy and to discuss its usefulness and rationale | Definition: as below  
Clinical features: assessment of leak was based on clinical signs (peritonitis, pyrexia and sepsis), morbidity and roentgenograms  
Investigations: all patients had a Gastrografin swallow 1 week postoperatively |

continued
<table>
<thead>
<tr>
<th>Study</th>
<th>Type of surgery, aim of study</th>
<th>Definition, clinical features, investigations</th>
</tr>
</thead>
</table>
| Hansen et al., 1996, Germany | Surgery: left-sided colon resection (n = 615) | **Definition:** clinical leaks were defined as below  
**Clinical features:** septic complications (abscess, leak (see below) or peritonitis)  
**Investigations:** a radiological examination of the anastomosis was performed only in clinically suspected cases of anastomotic leak (intra-abdominal abscess, postoperative peritonitis, faecal-stained discharge from drains). All anastomoses were tested intraoperatively by transanal instillation of povidone iodine solution |
| Hansson et al., 1997, Sweden | Surgery: gastro-oesophageal resection (n = 53) | **Definition:** none  
**Clinical features:** not described  
**Investigations:** a water-soluble radiographic swallow examination was performed in the majority of patients 3–28 days (median 7 days) after surgery |
| Hardy et al., 1996, Australia | Surgery: orthopotic liver transplants for end-stage liver disease (n = 129) | **Definition:** biliary tract complications were defined as any morbidity related to the biliary reconstruction. They were suspected on clinical, biochemical, microbiological or liver biopsy evidence  
**Clinical features:** as above  
**Investigations:** biliary stenoses were defined as radiological evidence of bile duct narrowing by cholangiography (stent tube, endoscopic retrograde cannulation) or percutaneous transhepatic cholangiography. All cholangiograms were reviewed separately by two independent observers. (Biliary leak occurred in 7 patients on removal of their biliary stent 3 months post-transplant) |
| Hida et al., 1996, Japan       | Surgery: low anterior resection (n = 43) | **Definition:** none  
**Clinical features:** not described  
**Investigations:** where a diverting colostomy was used, this was closed 2 months later when there was no evidence of anastomotic leak on a water-soluble contrast enema |
| Honkoop et al., 1996, Netherlands | Surgery: oesophagectomy and oesophagogastronomy (n = 269) | **Definition:** none  
**Clinical features:** not described  
**Investigations:** all but 3 patients (1%) underwent a radiographic swallow examination with water-soluble contrast medium between the days 7 and 10 postoperatively to detect anastomotic leak |
| Howard, 1997, USA | Surgery: pancreatojejunostomy (n = 56) | **Definition:** none  
**Clinical features:** not described  
**Investigations:** serum amylase levels were determined at 2-day intervals for the first 8–12 days postoperatively in most patients. Amylase levels were determined on the same day on the peritoneal (drain) and pancreatic (tube) fluid at least once between days 7 and 10 postoperatively |
<table>
<thead>
<tr>
<th>Study</th>
<th>Type of surgery, aim of study</th>
<th>Definition, clinical features, investigations</th>
</tr>
</thead>
<tbody>
<tr>
<td>Hrung et al., 1998.</td>
<td>Surgery: ileoanal pouch (n = 40)</td>
<td><strong>Definition:</strong> none</td>
</tr>
</tbody>
</table>
|                       | **Aim:** to assess the role of contrast enemas for the evaluation of postoperative leaks in asymptomatic and symptomatic patients after the first stage of restorative proctocolectomy | **Clinical features:** two patients with abdominal pain and fever underwent contrast enemas, and one patient with abdominal pain and distension underwent contrast enema  
**Investigations:** clinical leaks were investigated with water-contrast and/or barium enema to assess the anatomy and function of the pouch and rule out a subclinical leak. Routine imaging (late) was done at 8–12 weeks |
| Hulten, 1994.          | Surgery: restorative proctectomy (n = 307)                                                     | **Definition:** none                                                                 |
|                       | **Aim:** to analyse the complications and eventual outcome in 307 patients undergoing a pelvic pouch reconstruction | **Clinical features:** when clinical signs of leak or sepsis or radiological leak were demonstrated, ileostomy closure was delayed until endoscopic and radiological examinations were satisfactory. No description was given of clinical signs of leak  
**Investigations:** closure of the ileostomy after 6–8 weeks was always preceded by endoscopy and radiological examination of the pouch (gastrography introduced from above via the polyethylene irrigation tube) to ensure an intact anastomosis and suture lines |
| Iversen et al., 1999.  | Surgery: lower anterior resection of colon (n = 161)                                           | **Definition:** none                                                                 |
|                       | **Aim:** to study whether patients developing anastomotic dehiscence after colorectal resections for colorectal cancer had signs of an altered haemostatic balance in the systemic circulation, preoperatively and postoperatively, causing an impaired healing process | **Clinical features:** patients who developed overt clinical symptoms during the early postoperative period underwent soluble contrast enema. No description was given of clinical signs or symptoms  
**Investigations:** soluble contrast enema was used to confirm clinical suspicion. Blood testing for haemostatic markers was undertaken on patients with anastomotic leak and 17 age- and sex-matched control patients from the same cohort. Samples were taken for analysis of: prothrombin fragment 1 and 2, enzyme–inhibitor complexes produced during activation (thrombin–antithrombin complexes) and soluble fibrin |
| Jacobi et al., 1998.   | Surgery: oesophagectomy (n = 33)                                                              | **Definition:** none                                                                 |
|                       | **Aim:** to evaluate the perioperative changes of the submucosal P02 in the human stomach during oesophagectomy and to analyse the influence of P02 on anastomotic healing of the oesophagogastrostomy | **Clinical features:** anastomotic integrity was assessed clinically, but no description was given  
**Investigations:** anastomotic integrity was assessed clinically and by radiological evaluation using a Gastrografin swallow examination on day 7 postoperatively in every patient. All patients were followed up at 3-month intervals, with gastroscopy and radiological examination by Gastrografin swallow to detect the development of late fibrotic stenosis |
| Junger et al., 1996.   | Surgery: colorectal surgery (n = 22)                                                           | **Definition:** none                                                                 |
|                       | **Aim:** to determine whether the measurement of endotoxins or lipopolysaccharides in peritoneal fluid has any value to assess colorectal anastomotic integrity in the early postoperative period | **Clinical features:** clinical signs of anastomotic leak were reported as faecal secretion into the drain and diffuse peritonitis with intra-abdominal abscess. Leak preoperatively was not clearly defined  
**Investigations:** lipopolysaccharide concentrations in the drainage fluid. The peritoneal fluid (drainage fluid) was collected daily. See p. 68 for a full description |
| Kapoor et al., 1996.   | Surgery: intrahepatic segment III cholangiojejunostomy (n = 41)                                | **Definition:** none                                                                 |
|                       | **Aim:** to present the experience with intrahepatic segment III cholangiojejunostomy for palliation of symptoms in patients with advanced gallbladder cancer in a large single-centre study | **Clinical features:** not described  
**Investigations:** isotope hepatobiliary scanning was performed in a subset of patients with recurrent jaundice or cholangitis |
### TABLE 25 contd  Studies of anastomotic leak

<table>
<thead>
<tr>
<th>Study</th>
<th>Type of surgery, aim of study</th>
<th>Definition, clinical features, investigations</th>
</tr>
</thead>
</table>
| Karanjia et al., 1994, 274 UK | Surgery: stapled low anterior resection (n = 276)  
Aim: a study of anterior resection for rectal carcinoma | **Definition**: anastomotic leak was documented if it was shown in an asymptomatic patient on contrast enema, if it produced unequivocal clinical evidence such as peritonitis requiring laparotomy, or if clinical suspicion led to a contrast enema that confirmed a leak. Leak was classified as: minor (produced no clinically significant disturbance, but was detected by contrast enema) and major (produced clinically significant effects such as peritonitis or discharge of a pelvic collection)  
**Clinical features**: as above  
**Investigations**: contrast enema |
| Kartheuser et al., 1996, 298 France | Surgery: ileal pouch–anal anastomosis in familial adenomatous polyposis (n = 171)  
Aim: a prospective study to determine the incidence of complications and the functional results after ileal pouch–anal anastomosis with mucosectomy in a series of 171 patients with familial adenomatous polyposis, and to compare these functional results with those after ileorectal anastomosis (published previously) | **Definition**: none  
**Clinical features**: not described  
**Investigations**: all patients had a temporary diverting loop ileostomy, which was closed after roentgenological examination and routine rectal digital and the functional results after ileal pouch–anal anastomosis |
| Kayshara et al., 1995, 197 Japan | Surgery: pancreaticoduodenectomy (n = 150)  
Aim: to describe a new and simple technique of facilitating continuous intra-abdominal suction drainage following pancreaticoduodenectomy, and to discuss its effectiveness in reducing the incidence of post-operative complications | **Definition**: a diagnosis of dehiscence of the anastomosis was made in accordance with any of the following criteria: drainage of bile or enteric juice from the drain, the detection of enteric bacteria in the drainage fluid, radiographic confirmation of dehiscence of pancreatic ductography, or an amylase level in the drainage fluid of > 3 times the serum amylase level  
**Clinical features**: not described  
**Investigations**: radiographic confirmation of dehiscence of the pancreatic ductography |
| Kelly et al., 1994, 273 UK | Surgery: ileoanal pouch surgery (n = 85)  
Aim: to describe a simple and safe technique for examining the ileal pouch prior to ileostomy closure | **Definition**: anastomotic dehiscence was defined by extravasation of contrast at pouchography  
**Clinical features**: not described  
**Investigations**: pouchography was routinely performed 5–6 weeks after the first-stage procedure prior to closure of the covering ileostomy |
| Kessler et al., 1993, 274 Germany | Surgery: colorectal surgery (n = 1115)  
Aim: a prospective clinical–pathological observation study to analyse the present state of rectum carcinoma surgery and to seek indicators of clinical relevance | **Definition**: none  
**Clinical features**: clinical manifestations were stool fistulas, local abscess, persisting purulent secretion from drainage and peritonitis  
**Investigations**: testing of anastomoses was by methylene blue. No routine radiological controls were carried out |
| Kockerling et al., 1999, 285 Germany, Austria and Switzerland | Surgery: laparoscopic or laparoscopically assisted colorectal surgery (n = 1143)  
Aim: to present the results of a multicentre study in the form of data obtained prospectively from more than 1000 consecutive patients undergoing laparoscopic procedures. To investigate the safety of laparoscopic colorectal surgery on the basis of the anastomotic insufficiency rates in various sections of the bowel and to compare these with those seen in open colorectal surgery | **Definition**: despite being a major multicentre trial in centres from three countries, no clear definition was reported. Clinical anastomotic leaks were presented as the number and rate per procedure  
**Clinical features**: clinical anastomotic leaks were reported but not described  
**Investigations**: no details given |

*continued*
### TABLE 25 contd  Studies of anastomotic leak

<table>
<thead>
<tr>
<th>Study</th>
<th>Type of surgery, aim of study</th>
<th>Definition, clinical features, investigations</th>
</tr>
</thead>
</table>
| Kracht et al., 1993, France, French Association for Surgical Research | Surgery: right hemicolectomy (n = 440)  
**Aim:** to compare the incidence of anastomotic leak between four techniques of sutured and one type of stapled ileocolonic anastomosis after right colectomy for carcinoma in a prospective multicentre (29 centres) randomised study | **Definition:** the main end-point was anastomotic failure, which included all clinical leaks (assessed by the presence of an intra-abdominal abscess, postoperative peritonitis or faecal-stained discharge through the incision or drains) and any extravasation seen on a routine sodium diatrizoate enema performed between days 8 and 10 postoperatively. The presence of leak during re-laparotomy or necropsy was also noted.  
**Clinical features:** as above  
**Investigations:** routine sodium diatrizoate enema between days 8 and 10 postoperatively. The radiologist was blinded to the type of anastomosis performed |
| Kusano et al., 1997, Japan | Surgery: oesophagectomy (n = 115)  
**Aim:** to assess the impact of haemodynamic and oxygen transport variables on postoperative complications in patients undergoing oesophagectomy for carcinoma of the thoracic oesophagus | **Definition:** anastomotic leak was documented by means of clinical signs or radiography findings at days 3–12 postoperatively.  
**Clinical features:** not described  
**Investigations:** not fully described |
| Kuwano et al., 1993, Japan | Surgery: thoracic oesophagectomy (n = 343)  
**Aim:** to present the clinical results of patients with intrathoracic anastomotic leak who were treated with mediastinal drainage | **Definition:** as below  
**Clinical features:** a drain tube was fixed to the anastomosis and continuous suction applied postoperatively. If there was no purulent drainage, the drain was clamped to allow movement. In the presence of anastomotic leak and a purulent discharge through the drain, suction was continued until the amount of discharge was reduced.  
**Investigations:** If there was purulent drainage through the drain at 2 weeks postoperatively a radiological study of Gastrografin swallow and an infusion of Urografin through the drain were done and the abscess cavity of the posterior mediastinum and condition of leak evaluated |
| Law et al., 1997, Hong Kong | Surgery: oesophagectomy (n = 122)  
**Aim:** a prospective randomised trial comparing a single-layer continuous hand-sewn method with the use of a circular stapler in a uniform population undergoing Lewis–Tanner oesophagectomy | **Definition:** none  
**Clinical features:** diagnosis of anastomotic leak was based on clinical (septic complications) and on radiological evidence  
**Investigations:** all patients were assessed using a Gastrografin contrast study on day 7 postoperatively. Endoscopy was also done |
| Lowy et al., 1997, USA | Surgery: pancreaticoduodenectomy (n = 120)  
**Aim:** a prospective randomised study to test the hypothesis that peri-operative octreotide decreases the incidence of pancreatic anastomotic leak after pancreaticoduodenectomy for malignant disease | **Definition:** a clinical pancreatic anastomotic leak was defined as the drainage of amylase-rich fluid (> 2.5 times the upper limit of normal for serum amylase) in association with fever (> 38°C), leucocytosis (white blood cell count > 10,000/l), sepsis (haemodynamic instability requiring transfer to the intensive care unit) or the need for percutaneous drainage of an amylase-rich fluid collection. A biochemical pancreatic leak was defined as an elevated level of amylase (> 2.5 times the upper limit of normal for serum amylase) in the drain fluid, on or after postoperative day 3, that was asymptomatic and resolved spontaneously.  
**Clinical features:** as above  
**Investigations:** the amylase content of the fluid from the medial abdominal drain was measured and recorded on day 5 postoperatively. In the absence of a clinical or anastomotic leak, drains were removed before day 8 postoperatively |

*continued*
TABLE 25 contd Studies of anastomotic leak

<table>
<thead>
<tr>
<th>Study</th>
<th>Type of surgery, aim of study</th>
<th>Definition, clinical features, investigations</th>
</tr>
</thead>
</table>
| Machens et al., 1996, Germany | Surgery: oesophagectomy (n = 12) | **Definition**: as below  
**Clinical features**: when a cervical leak was clinically suspected as a result of local inflammation at the anastomotic site or when air or saliva was found in the cervical drain bag, the neck wound was opened widely and explored. Cervical leaks confirmed on exploration of the neck wound were termed major, and minor leaks were identified only in routine contrast studies  
**Investigations**: a Gastrografin swallow was performed on day 7 postoperatively to check for cervical leak. Major leaks manifested by local inflammation at the anastomotic site. Median drain amylase analysis was also done |
| Mann et al., 1996, Germany | Surgery: colorectal resection (n = 370) | **Definition**: as below  
**Clinical features**: anastomotic leak was suspected if the patient showed fever, leucocytosis, persistent ileus, bleeding or discharge, or any other sign of intra-abdominal or pelvic abscess  
**Investigations**: routine contrast enema was not performed. In patients with the above clinical features, a water-soluble contrast enema was carried out immediately |
| Mansour et al., 1997, USA | Surgery: oesophagectomy (n = 131) | **Definition**: none  
**Clinical features**: not described  
**Investigations**: a Gastrografin oesophagogram, followed by barium swallow if the Gastrografin study showed no leak, was performed between days 6 and 9 postoperatively. Patients with small cervical leaks were allowed to begin oral fluids |
| Matsusue et al., 1998, Japan | Surgery: pancreaticoduodenectomy (n = 100) | **Definition**: protracted healing of the pancreaticojejunostomy was divided into peripancreatic sepsis and pancreatic fistula. Peripancreatic sepsis was defined as prolonged suppurative discharge of < 50 ml/day with a low amylase level (< 1000 IU) from a drain beneath the pancreaticojejunostomy for more than 1 week. Pancreatic fistula was defined as prolonged discharge of > 50 ml/day with a high amylase level (> 1000 IU) from the drain for more than 1 week  
**Clinical features**: as above  
**Investigations**: as above |
| Merad et al., 1998, France | Surgery: colorectal resection (n = 712) | **Definition**: none  
**Clinical features**: anastomotic leak was diagnosed by the egress of faecal fluid through drains, by repeat operation or at autopsy (performed routinely for all patients who died during their hospital stay), or by sodium diatrizoate enema, which was performed routinely near day 7 for asymptomatic patients  
**Investigations**: leak was detected by routine water-soluble (sodium diatrizoate) radiograms alone. Anastomotic leak was defined as all images other than those showing a perfectly regular and uniform calibre at the level of the anastomosis |
| Merad et al., 1999, France | Surgery: colorectal resection (n = 494) | **Definition**: radiological leak was defined as all images except those showing a perfectly regular and uniform calibre at the level of the anastomosis  
**Clinical features**: anastomotic leak was diagnosed either by faeces exiting through drains, by repeat operation or at autopsy (performed routinely for all patients who died during the hospital stay), or by sodium diatrizoate enema, which was performed routinely on or near day 8 for asymptomatic patients  
**Investigations**: sodium diatrizoate enema performed routinely on or near day 8 for asymptomatic patients |
<table>
<thead>
<tr>
<th>Study</th>
<th>Type of surgery, aim of study</th>
<th>Definition, clinical features, investigations</th>
</tr>
</thead>
<tbody>
<tr>
<td>Miller and Moritz, 1996, Austria</td>
<td>Surgery: low anterior resection of rectum (n = 103)</td>
<td>Definition: as below</td>
</tr>
<tr>
<td></td>
<td>Aim: to report on the experience with stapled anastomosis in patients with potentially curative resection of rectal carcinoma</td>
<td>Clinical features: clinical leaks were said to occur if pus appeared per anum, if faeces appeared in the drain, or if pelvic abscess developed</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Investigations: a water-soluble contrast enema was undertaken on day 10 postoperatively. Any leak of contrast material was interpreted as a radiological leak</td>
</tr>
<tr>
<td>Miller et al., 1996, Austria</td>
<td>Surgery: low anterior resection of the rectum (n = 42)</td>
<td>Definition: as below</td>
</tr>
<tr>
<td></td>
<td>Aim: to investigate if there is a difference in lysozyme activity in the drained liquid secretion between the single-staple technique and the double-staple technique in anterior resection of rectal carcinoma, and if there is a difference between these two techniques with regard to anastomotic dehiscence</td>
<td>Clinical features: clinical leaks were said to occur if pus appeared per anum, if faeces appeared in the drain, or if pelvic abscess developed</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Investigations: a water-soluble contrast enema was undertaken on day 10 postoperatively. Any leak of contrast material was interpreted as a radiological leak</td>
</tr>
<tr>
<td>Moore et al., 1996, Australia</td>
<td>Surgery: colorectal surgery, rectal resection (n = 826)</td>
<td>Definition: as below</td>
</tr>
<tr>
<td></td>
<td>Aim: to describe the experience with double-stapled anastomosis and to compare the morbidity, mortality and cancer-related outcomes associated with single- and double-stapled anastomoses performed in a single institution</td>
<td>Clinical features: clinically significant (generalised) anastomotic leak was defined as that necessitating urgent laparotomy. Subclinical (localised) leak was managed conservatively. No description was given of what was meant by ‘subclinical’ or how such leaks were detected</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Investigations: radiological examination of the anastomosis was performed only when anastomotic leak was suspected on clinical grounds, or prior to closure of a proximal diverting stoma. Radiological examination of the anastomosis was not routinely performed</td>
</tr>
<tr>
<td>Nagakawa et al., 1997, Japan</td>
<td>Surgery: pancreatoduodenectomy (n = 64)</td>
<td>Definition: leak of pancreatic juice was defined as when the amylase level in the drainage from the anastomotic site between the pancreas and the intestine was ≥ 1000 IU/I. Suture insufficiency was defined as when the intestinal contents, particularly bile, were contaminated</td>
</tr>
<tr>
<td></td>
<td>Aim: modifications of techniques for pancreaticojejunostomy are described, with an assessment of the complication rate for various anastomotic techniques</td>
<td>Clinical features: not assessed</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Investigations: as above</td>
</tr>
<tr>
<td>Nambirajan et al., 1998, UK</td>
<td>Surgery: primary repair of oesophageal atresia (n = 49)</td>
<td>Definition: anastomotic leaks were defined as incidental (small radiological leak, no clinical symptoms), minor (saliva in chest drain, but clinically well) and major (mediastinitis or abscess, pneumothorax, empyema, radiologically confirmed major oesophageal disruption)</td>
</tr>
<tr>
<td></td>
<td>Aim: to determine the incidence of postoperative leaks as seen on an early oesophagram and the relationship between an anastomotic leak and subsequent stricture formation</td>
<td>Clinical features: as above</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Investigations: a postoperative water-soluble contrast oesophagram was obtained on days 5 to 7 to determine the presence or absence of leak before oral feeds were commenced</td>
</tr>
<tr>
<td>Norris et al., 1999, Australia</td>
<td>Surgery: laparotomy for Crohn's disease (n = 156)</td>
<td>Definition: none</td>
</tr>
<tr>
<td></td>
<td>Aim: the primary aim was to establish, using data from a prospective database, whether patients with Crohn's disease who required a laparotomy at an older age have comparable morbidity and mortality</td>
<td>Clinical features: not described</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Investigations: all postoperative anastomotic leaks were confirmed by imaging or re-operation</td>
</tr>
</tbody>
</table>

continued
### TABLE 25 contd  
**Studies of anastomotic leak**

<table>
<thead>
<tr>
<th>Study</th>
<th>Type of surgery, aim of study</th>
<th>Definition, clinical features, investigations</th>
</tr>
</thead>
</table>
| Obertop and Bosscha, 1994, Netherlands | **Surgery:** oesophageal resection and reconstruction (n = 10)  
**Aim:** to report the results of a radical surgical approach in patients with life-threatening mediastinitis resulting from anastomotic disruption with regard to survival, morbidity and digestive tract reconstruction  
**Definition:** anastomotic disruption and mediastinitis were diagnosed by clinical signs of sepsicaemia in all patients, and leak was confirmed by surgery in all patients  
**Clinical features:** as above  
**Investigations:** confirmed by surgery. Contrast oesophagography and CT were done only in a subset of patients, not routinely | |
| O’Rourke et al., 1995, Australia | **Surgery:** oesophagogastrectomy (n = 77)  
**Aim:** to review the current peri-operative and long-term outcome of all patients undergoing surgery for oesophageal cancer in a specialised teaching hospital unit, with careful attention paid to patient selection and preparation, specialised anaesthesia and intensive postoperative care  
**Definition:** none  
**Clinical features:** not defined  
**Investigations:** Gastrografin swallow on day 7 to 10 postoperatively before oral feeding was begun, and nasogastric drainage until anastomosis was checked radiographically | |
| Pakkastie et al., 1997, Finland | **Surgery:** low anterior colorectal resection for rectal neoplasms (n = 38)  
**Aim:** to evaluate the value of a covering colostomy on the anastomotic leak rate and on the outcome of patients after low anterior resection  
**Definition:** radiological leak was defined as any sign of the contrast medium outside the bowel wall  
**Clinical features:** all patients were followed for clinical signs of leak (faecal discharge from the wound or drain, pelvic sepsis, postoperative fever or sepsicaemia)  
**Investigations:** a water-soluble contrast enema was given between day 7 and 10 postoperatively, or before this time to confirm the presence of a leak if there was any suspicion of it | |
| Petersen et al., 1998, Germany | **Surgery:** anterior resection of the colon (n = 467)  
**Aim:** to determine the effect of anastomotic leak on tumour recurrence and survival. Of special interest was the comparison between patients who developed anastomotic leak and those who did not, with reference to well-known prognostic factors  
**Definition:** a clinical leak was defined as an anastomotic dehiscence, which was confirmed by either the appearance of faeces from the wound or the drains, or by the development of a colo- or rectocutaneous fistula. Gastrografin enema for detection of subclinical leak was not performed routinely  
**Clinical features:** as above  
**Investigations:** Gastrografin enema for the detection of subclinical leak was not performed routinely | |
| Pickleman et al., 1999, USA | **Surgery:** oesophagogastrectomy, total or partial gastrectomy, enterectomy and partial or subtotal colectomy (n = 2842)  
**Aim:** a case–control study to analyse the variables in patients undergoing gastrointestinal anastomoses at all levels of the gastrointestinal tract, and to study the impact of anastomotic disruption, with a focus on diagnosis, treatment and outcomes. Data from a prospectively collected data set were used  
**Definition:** oesophagogastrectomy was diagnosed by the appearance of gastrointestinal contents in the thoracostomy tube drainage output (subset of patients); CT drainage of loculated fluid (n = 1). Total gastrectomy was diagnosed by Gastrografin studies of the upper gastrointestinal tract (n = 3) and CT scan (n = 1). Partial gastrectomy was diagnosed clinical grounds only (n = 2) and Gastrografin contrast studies (n = 2). Enterectomy was diagnosed on clinical grounds only (n = 4), CT scan (n = 4) and abdominal X-ray (n = 1). Partial colectomy was diagnosed on clinical grounds only (n = 3), abdominal X-ray (n = 2), CT scan (n = 13) and Gastrografin enema (n = 3). Subtotal colectomy was diagnosed on clinical grounds only (n = 1), abdominal X-ray (n = 1), CT scan (n = 2) and lower gastrointestinal tract contrast studies (n = 3)  
**Clinical features:** see below  
**Investigations:** oesophagogastrectomy was diagnosed by the appearance of gastrointestinal contents in the thoracostomy tube drainage output (subset of patients); CT drainage of loculated fluid (n = 1). Total gastrectomy was diagnosed by Gastrografin studies of the upper gastrointestinal tract (n = 3) and CT scan (n = 1). Partial gastrectomy was diagnosed clinical grounds only (n = 2) and Gastrografin contrast studies (n = 2). Enterectomy was diagnosed on clinical grounds only (n = 4), CT scan (n = 4) and abdominal X-ray (n = 1). Partial colectomy was diagnosed on clinical grounds only (n = 3), abdominal X-ray (n = 2), CT scan (n = 13) and Gastrografin enema (n = 3). Subtotal colectomy was diagnosed on clinical grounds only (n = 1), abdominal X-ray (n = 1), CT scan (n = 2) and lower gastrointestinal tract contrast studies (n = 3) | |

*continued*
<table>
<thead>
<tr>
<th>Study</th>
<th>Type of surgery, aim of study</th>
<th>Definition, clinical features, investigations</th>
</tr>
</thead>
<tbody>
<tr>
<td>Pol et al., 1997,220</td>
<td>Surgery: total gastrectomy (n = 176)</td>
<td>Definition: none</td>
</tr>
<tr>
<td></td>
<td>Aim: to evaluate the feasibility and reliability of stapled oesophagojejunostomy in a prospective series of 176 consecutive total gastrectomies performed by the abdominal approach only, regardless of the indications</td>
<td>Clinical features: assessed clinically, but no details given</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Investigations: anastomotic integrity was checked in week 2 postoperatively using a water-soluble contrast medium. Leak was also revealed by a subphrenic abscess and during re-operation for peritonitis and mediastinitis</td>
</tr>
<tr>
<td>Redmond et al., 1993,259</td>
<td>Surgery: low anterior resection of the colon (n = 111)</td>
<td>Definition: radiological leak was defined as radio-opaque dye clearly visible in the perianastomotic space, with or without clinical evidence of leak</td>
</tr>
<tr>
<td></td>
<td>Aim: to evaluate the use of double-stapled anastomosis in low anterior resection</td>
<td>Clinical features: clinical leak was defined as a breakdown in the double-spaced anastomosis, resulting in clinical signs of peritonitis or septic shock such that the patient required laparotomy</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Investigations: Gastrografin enema was administered on day 10 to 12 postoperatively to assess the radiological integrity of the anastomosis. Results were presented as clinical leak and radiological leak</td>
</tr>
<tr>
<td>Reissman et al., 1995,253</td>
<td>Surgery: pancreaticoduodenectomy (n = 35)</td>
<td>Definition: pancreaticojejunoanastomosis anastomotic leak was defined as the recovery of &gt; 40 ml/day of amylase-rich fluid (&gt; 10 times normal plasma level) from the peripancreatic drains for &gt; 7 days, or a radiologically demonstrable leak</td>
</tr>
<tr>
<td></td>
<td>Aim: to prospectively assess a technique that avoids pancreaticojejunoanastomosis (i.e. ligation of the pancreatic duct and placement of drains to create a temporary controlled pancreaticocutaneous fistula)</td>
<td>Clinical features: not described</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Investigations: no description was given of the radiological test</td>
</tr>
<tr>
<td>Richard et al., 1997,255</td>
<td>Surgery: pelvic pouch procedure (n = 753)</td>
<td>Definition: none</td>
</tr>
<tr>
<td></td>
<td>Aim: a prospective study to determine the effect of prior perianal disease on the short- and long-term outcomes of pelvic pouch patients</td>
<td>Clinical features: not described</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Investigations: all ileoanal leaks were confirmed by either radio-opaque contrast enema or by CT with contrast</td>
</tr>
<tr>
<td>Rodella et al., 1998,214</td>
<td>Surgery: oesophago gastric surgery (n = 7)</td>
<td>Definition: none</td>
</tr>
<tr>
<td></td>
<td>Aim: to report the first experience in the treatment of anastomotic leaks in oesophago gastric surgery with endoscopic clipping</td>
<td>Clinical features: not described</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Investigations: in all cases diagnosis was made with Gastrografin swallow and endoscopy</td>
</tr>
<tr>
<td>Roder et al., 1999,215</td>
<td>Surgery: pancreatoduodenectomy (n = 85)</td>
<td>Definition: as below</td>
</tr>
<tr>
<td></td>
<td>Aim: a prospective study to compare the morbidity and mortality rates of stented versus non-stent pancreaticojejunoanastomosis after partial pancreatoduodenectomy</td>
<td>Clinical features: clinical symptoms (fever, elevated leucocyte count, sepsis), radiographically documented leaks, or a fluid collection in the surgical drain adjacent to the pancreaticojejunoanastomosis of &gt; 50 cm² during the entire course after surgery, with an amylase content exceeding triple the serum amylase concentration, were considered indicative of a fistula of the pancreaticojejunoanastomosis</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Investigations: as above. Surgical drains were connected to a closed drainage system to detect and drain potential leaks after surgery</td>
</tr>
</tbody>
</table>

continued
### TABLE 25 contd  Studies of anastomotic leak

<table>
<thead>
<tr>
<th>Study</th>
<th>Type of surgery, aim of study</th>
<th>Definition, clinical features, investigations</th>
</tr>
</thead>
<tbody>
<tr>
<td>Sagar et al., 1995, 281 UK</td>
<td>Surgery: colorectal resection (n = 284)</td>
<td><strong>Definition:</strong> as below &lt;br&gt; <strong>Clinical features:</strong> a clinically significant anastomotic leak was defined as discharge of faeces from the drain site or presence of an abscess in close proximity to the anastomosis, and localised or generalised peritonitis with tenderness, fever and leucocytosis &lt;br&gt; <strong>Investigations:</strong> patients underwent water-soluble contrast enema (Gastrografin) on day 5 to 7 postoperatively to test the integrity of the anastomosis. The presence of a leak indicated by the Gastrografin enema in a patient who had no clinical evidence of problems with the anastomosis was recorded as a radiological leak</td>
</tr>
<tr>
<td>Santos et al., 1994, 282 Brazil</td>
<td>Surgery: colorectal resection (n = 232)</td>
<td><strong>Definition:</strong> as below &lt;br&gt; <strong>Clinical features:</strong> investigation was undertaken in the presence of fever, abdominal or sacral pain, tenesmus, or clinical signs of localised or generalised peritonitis &lt;br&gt; <strong>Investigations:</strong> patients without clinical symptoms were not submitted for investigation. Anastomotic dehiscence was confirmed on clinical, radiographic or intraoperative examination</td>
</tr>
<tr>
<td>Schardey et al., 1997, 283 Germany</td>
<td>Surgery: total gastrectomy (n = 205)</td>
<td><strong>Definition:</strong> anastomotic insufficiency and anastomotic leak were defined as a complete intestinal wall defect at the anastomotic suture line, as detected by radiological contrast medium study or by a positive colour test &lt;br&gt; <strong>Clinical features:</strong> not described. Routine clinical chemistry and haematological tests were not described &lt;br&gt; <strong>Investigations:</strong> a Gastrografin swallow was carried out when oesophagointestinal leak was suspected, or routinely on day 7 postoperatively. If the patient was unable to cooperate or was too ill, a bedside test with Gastrografin or indigo carmine blue in the intensive treatment unit or endoscopic proof with radiological Gastrografin documentation, or an intraoperative test with indigo carmine blue was carried out during emergency surgery to detect or exclude oesophagointestinal leak. Independent radiologists not involved in the study performed and interpreted the tests</td>
</tr>
<tr>
<td>Schilling et al., 1996, 284 Switzerland</td>
<td>Surgery: oesophagectomy (n = 11)</td>
<td><strong>Definition:</strong> none &lt;br&gt; <strong>Clinical features:</strong> not defined &lt;br&gt; <strong>Investigations:</strong> no patient had anastomotic necroses or leak as demonstrated by postoperative Gastrografin X-ray swallows</td>
</tr>
<tr>
<td>Schilling et al., 1997, 285 Switzerland</td>
<td>Surgery: oesophagectomy (n = 35)</td>
<td><strong>Definition:</strong> none &lt;br&gt; <strong>Clinical features:</strong> not described. Radiological and clinical evidence of anastomotic leak &lt;br&gt; <strong>Investigations:</strong> Gastrografin swallow radiographs taken of all patients between day 5 and 10 postoperatively and oral feeding was then started</td>
</tr>
<tr>
<td>Slim et al., 1995, 286 France</td>
<td>Surgery: laparoscopic and laparoscopically assisted colorectal resection (n = 65)</td>
<td><strong>Definition:</strong> as below &lt;br&gt; <strong>Clinical features:</strong> clinical anastomotic leak was denoted by faecal matter or pus in the drainage discharge, purulent discharge per anum or intra-abdominal abscess, or at re-operation for peritonitis &lt;br&gt; <strong>Investigations:</strong> a postoperative water-soluble enema was administered only in the case of suspected leak, not routinely</td>
</tr>
</tbody>
</table>

*RCT, randomised controlled trial*
<table>
<thead>
<tr>
<th>Study</th>
<th>Type of surgery, aim of study</th>
<th>Definition, clinical features, investigations</th>
</tr>
</thead>
</table>
| Solomon et al., 1995, Canada | Surgery: ileoanal pouch (n = 16)  
Aim: to prospectively evaluate early results of endoluminal transpouch ultrasonography in patients with symptoms that could be related to peripouch sepsis with or without an anastomotic leak | Definition: none  
Clinical features: not described  
Investigations: different methods were reported, including pouchography, endoluminal transpouch ultrasonography and CT |
| Spiliopoulos et al., 1998, Switzerland | Surgery: intrathoracic oesophagectomy (n = 43)  
Aim: to describe a useful and simple to perform adjunct to the standard Ivor–Lewis procedure that may reduce the consequences of potential anastomotic leak | Definition: none  
Clinical features: not described  
Investigations: Gastrografin swallow performed within 10 days postoperatively |
| Stewart et al., 1998, Australia | Surgery: colorectal resection without stoma formation (n = 88)  
Aim: a prospective randomised trial to evaluate further the potential benefits and risks of early feeding compared to traditional management in patients undergoing open elective colorectal resection | Definition: none  
Clinical features: one patient had an anastomotic leak confined to a faecal discharge from the drain tube  
Investigations: not described |
| Sugerman and Newsome, 1994, USA | Surgery: ileoanal pouch surgery (n = 75)  
Aim: to report on the experience of patients undergoing a stapled ileoanal anastomosis without a temporary ileostomy | Definition: none  
Clinical features: there are reports in the results of patients having faecal drainage in their Jackson–Pratt drains  
Investigations: those with symptoms underwent water-soluble contrast enemas |
| Svanes et al., 1995, Norway | Surgery: oesophagectomy (n = 83)  
Aim: a prospective study to gain information about treatment policy for patients with cancer of the oesophagus and cardia, in particular about the extent of early and late postoperative morbidity, which contributes to a reduction in the quality of their remaining life | Definition: none  
Clinical features: not described  
Investigations: anastomoses were examined at days 8 to 10 postoperatively with water-soluble contrast to check for leaks |
| Swails et al., 1995, USA | Surgery: oesophagogastrectomy (n = 25)  
Aim: a prospective randomised study to assess the safety and efficacy of enteral feeding via jejunostomy placed at the time of elective oesophagogastrectomy | Definition: none  
Clinical features: not described  
Investigations: group 1 patients with a feeding jejunostomy tube placed at the time of surgery had oral feeding commenced after a contrast radiographical study (average time for feeding 8 ± 4 days). Group 2 patients who did not have a feeding jejunostomy tube received intravenous fluid replacement and underwent radiographical assessment on day 4 or 5 postoperatively |
| Tagart, 1986, UK | Surgery: colorectal resection (n = 220)  
Aim: to report a personal audit to assess whether restorative rectal resection is as safe as total rectal excision and whether it is as effective in the treatment of cancer | Definition: none  
Clinical features: not described  
Investigations: every anastomosis, excluding those in whom there was clinical evidence of a leak and non-survivors, was tested with a limited barium enema at the end of week 2 postoperatively (day 14) |
TABLE 25 contd Studies of anastomotic leak

<table>
<thead>
<tr>
<th>Study</th>
<th>Type of surgery, aim of study</th>
<th>Definition, clinical features, investigations</th>
</tr>
</thead>
</table>
| Thiede et al., 1998, 209 Germany and Austria | Surgery: upper and lower gastrointestinal surgery (n = 1360)  
Aim: a European prospective multicentre (n = 6) clinical trial to evaluate the applicability of the biofragmentable anastomosis ring as a routine anastomotic tool in teaching hospitals | Definition: none                                                                                               |
|                                            |                                                                                  | Clinical features: as below                                                                                   |
|                                            |                                                                                  | Investigations: all patients from one centre were submitted to radiological control of the anastomosis by water-soluble contrast medium at approximately day 8 postoperatively. All other patients were submitted to radiological examination if it was clinically indicated (retarded bowel function, fever, localised abdominal fluid) |
| Thomas et al., 1997, 234 France             | Surgery: colon interposition for oesophageal replacement (n = 60)  
Aim: to identify the current indications for use of the colon as an oesophageal substitute and to determine what effect several variables have on its long-term alimentary comfort | Definition: none                                                                                               |
|                                            |                                                                                  | Clinical features: not described                                                                               |
|                                            |                                                                                  | Investigations: the proximal oesophageojejunal anastomosis was checked routinely by radiography with water-soluble contrast medium on day 8 to 12 postoperatively, depending on the level of the anastomosis. All fistulas were counted, regardless of whether there was clinical evidence of leak |
| Thompson et al., 1996, 205 Australia        | Surgery: colorectal resection (n = 535)  
Aim: two publications related to the Australian Council for Health Care Standards proposal to test clinical indicators for colorectal resection; two audits conducted in the state of Victoria | Definition: defined as anastomotic leak detected clinically                                                   |
|                                            |                                                                                  | Clinical features: no description given of how leak was assessed                                                |
|                                            |                                                                                  | Investigations: as there was no requirement for contrast studies, radiologically detected leaks were not included in the analysis |
| Trentino et al., 1997, 217 Italy            | Surgery: oesophagectomy and cervical oesophagogastronomy (n = 39)  
Aim: to evaluate through early postoperative oesophagoscopy the morphological change of the anastomosis that could be related to benign anastomotic stenosis development | Definition: none                                                                                               |
|                                            |                                                                                  | Clinical features: not described                                                                               |
|                                            |                                                                                  | Investigations: all patients underwent Gastrografin swallow on day 8 or 9 postoperatively                      |
| van Lanschot et al., 1999, 228 Netherlands  | Surgery: oesophagectomy (n = 60)  
Aim: to determine the possible ‘price’ to be paid for the advantage of an extra-anatomical reconstruction by comparing the technical and functional results after prevertebral and retrosternal gastric tube reconstruction | Definition: not defined                                                                                       |
|                                            |                                                                                  | Clinical features: not described. Results were presented as clinical and radiological leaks                    |
|                                            |                                                                                  | Investigations: water-soluble contrast swallow examination on day 7 postoperatively                           |
| Vigneswaran et al., 1993, 229 USA           | Surgery: transhiatal oesophagectomy (n = 131)  
Aim: to evaluate transhiatal oesophagectomy as a cancer operation | Definition: none                                                                                               |
|                                            |                                                                                  | Clinical features: a subset of patients had leak that drained through the neck wound                           |
|                                            |                                                                                  | Investigations: roentgenographic oesophageal examination using a water-soluble, orally administered contrast agent was performed on (usually) day 7 postoperatively |
| Watson et al., 1999, 230 UK                 | Surgery: colorectal surgery (n = 477)  
Aim: a review of the experience of salvaged colorectal anastomoses using prospectively collected data | Definition: none                                                                                               |
|                                            |                                                                                  | Clinical features: not described                                                                               |
|                                            |                                                                                  | Investigations: dehiscence was confirmed or refuted by a water-soluble enema, as soon as clinical suspicion was raised |

continued
TABLE 25 contd Studies of anastomotic leak

<table>
<thead>
<tr>
<th>Study</th>
<th>Type of surgery, aim of study</th>
<th>Definition, clinical features, investigations</th>
</tr>
</thead>
<tbody>
<tr>
<td>Wexner et al., 1993, USA</td>
<td>Surgery: colorectal resection with loop ileostomy (n = 83)</td>
<td>Definition: none; Clinical features: not described; Investigations: all patients with anastomoses underwent contrast radiographical studies of the distal intestinal tract prior to closure of their ileostomy</td>
</tr>
<tr>
<td>Wheeler and Gilbert, 1999, UK</td>
<td>Surgery: colorectal resection (n = 102)</td>
<td>Definition: none; Clinical features: clinical and radiological features not described; Investigations: routine water-soluble contrast enema on day 8 postoperatively</td>
</tr>
<tr>
<td>Wu et al., 1995, Taiwan</td>
<td>Surgery: radical gastrectomy (n = 474)</td>
<td>Definition: none; Clinical features: not described; Investigations: the diagnosis of leak was confirmed by clinical data and radiological examination. Gastrografin was routinely given orally to ensure that the anastomoses were watertight before starting liquid food (time not specified)</td>
</tr>
<tr>
<td>Zieren et al., 1993, Germany</td>
<td>Surgery: oesophagectomy and cervical oesophagogastronomy (n = 107)</td>
<td>Definition: as below; Clinical features: not defined; Investigations: leak was defined as any radiographically demonstrable extravasation of water-soluble contrast medium at the site of anastomosis on day 7 postoperatively</td>
</tr>
<tr>
<td>Zilling et al., 1997, Sweden</td>
<td>Surgery: total gastrectomy (n = 174)</td>
<td>Definition: none; Clinical features: not described; Investigations: every anastomosis was checked radiologically for leak using water-soluble contrast on day 4 to 7 postoperatively</td>
</tr>
</tbody>
</table>
## TABLE 26 Prospective validation studies of anastomotic leak

<table>
<thead>
<tr>
<th>Study</th>
<th>Surgery, aim of study</th>
<th>Investigation, results</th>
</tr>
</thead>
<tbody>
<tr>
<td>Bischof et al., 1995, USA</td>
<td>Surgery: kidney–pancreas transplants (&lt;i&gt;n&lt;/i&gt; = 18)</td>
<td><strong>Investigation</strong>: use of air-contrast solution CT cystography. A leak from the duodenal segment was diagnosed if extravasal contrast material or air was demonstrated if the amount of pelvic or peripancreatic fluid on postvoiding scans was significantly increased. Two experienced blinded abdominal radiologists analysed the CT cystograms. <strong>Gold standard</strong>: findings at surgery, cystoscopy and multiple clinical follow-up examinations. <strong>Results</strong>: sensitivity 92%, specificity 100%, accuracy 96%</td>
</tr>
<tr>
<td>Goel et al., 1995, India</td>
<td>Surgery: oesophagectomy (&lt;i&gt;n&lt;/i&gt; = 25)</td>
<td><strong>Investigation</strong>: ingestion of drinking water (test feeding) after water-soluble contrast study, and neck site observed for leak. <strong>Gold standard</strong>: water-soluble (Gastrografin) study on day 5 postoperatively. <strong>Results</strong>: one patient did not tolerate contrast, two patients demonstrated leak on the contrast study but not on test feeding. Comparison table given; full details are given in the text of chapter 6</td>
</tr>
<tr>
<td>Kawano et al., 1995, Japan</td>
<td>Surgery: oesophageal resection (&lt;i&gt;n&lt;/i&gt; = 21)</td>
<td><strong>Investigation</strong>: platelet aggregability was measured preoperatively, and on days 1, 3 and 7 postoperatively. <strong>Gold standard</strong>: not applicable to this study. No details were given of how leak was detected clinically. There was a significant difference in platelet aggregability on day 1 postoperatively of patients with leak compared with those without. <strong>Results</strong>: the average values of platelet aggregability for those patients without leak were 81.2%, 70.4%, 80.1% and 81.8%, and those for patients with leak were 81.3%, 47.6%, 52.3% and 70.6%, preoperatively and on days 1, 3 and 7, respectively. There was a significant difference in platelet aggregability on day 1 postoperatively for patients with leak compared with those without</td>
</tr>
<tr>
<td>Miller et al., 1996, Austria</td>
<td>Surgery: low anterior resection of the rectum (&lt;i&gt;n&lt;/i&gt; = 42)</td>
<td><strong>Investigation</strong>: drained fluid secretions were collected for lysozyme determination from day 1 to 4 postoperatively. <strong>Gold standard</strong>: a water-soluble contrast enema was undertaken on the day 10 postoperatively. Any leak of contrast material was interpreted as a radiological leak. Clinical leaks were said to occur if pus appeared from the anus, if faeces appeared in the drain, or if pelvic abscess developed. <strong>Results</strong>: mean lysozyme activity was increased in those patients with clinically (18 mg/dl on day 1 postoperatively) and radiologically (15.3 mg/dl on day 1 postoperatively) detected dehiscence</td>
</tr>
<tr>
<td>Mouratidis et al., 1995, Australia</td>
<td>Surgery: pancreatic–renal allograft transplantation (&lt;i&gt;n&lt;/i&gt; = 1)</td>
<td><strong>Investigation</strong>: a single case report of anastomotic leak demonstrated by radionuclide cystography. <strong>Gold standard</strong>: not applicable. <strong>Results</strong>: anastomotic leak at the level of the transplanted duodenum was confirmed surgically and repaired</td>
</tr>
</tbody>
</table>

**continued**
## TABLE 26 contd  Prospective validation studies of anastomotic leak

<table>
<thead>
<tr>
<th>Study</th>
<th>Surgery, aim of study</th>
<th>Investigation, results</th>
</tr>
</thead>
<tbody>
<tr>
<td>Schilling et al., 1996, Switzerland</td>
<td>Surgery: oesophagectomy (n = 11)</td>
<td>Investigation: laser Doppler flowmetry was used to assess gastric microcirculatory changes during gastric tube formation</td>
</tr>
<tr>
<td></td>
<td><strong>Aim:</strong> to assess gastric microperfusion, by means of laser Doppler flowmetry, before, during and after gastric tube formation in patients undergoing oesophagectomy. A small non-randomised observational study</td>
<td><strong>Gold standard:</strong> postoperative Gastrografin X-ray</td>
</tr>
<tr>
<td></td>
<td></td>
<td><strong>Results:</strong> no leaks occurred. The authors suggest that the method may help prevent ischaemia-induced anastomotic breakdown</td>
</tr>
<tr>
<td>Simmen et al., 1994, Switzerland</td>
<td>Surgery: abdominal surgery (n = 55)</td>
<td>Investigation: analysis of pH, pO₂ and pCO₂ levels in peritoneal fluid monitored intraoperatively and postoperatively</td>
</tr>
<tr>
<td></td>
<td><strong>Aim:</strong> the pO₂ and pCO₂ of peritoneal fluid were monitored in addition to pH at laparotomy and during the postoperative follow-up period to evaluate the potential for rapid, simple and early detection of infectious complications following abdominal surgery</td>
<td><strong>Gold standard:</strong> anastomotic leak or abscess was confirmed radiologically</td>
</tr>
<tr>
<td></td>
<td></td>
<td><strong>Results:</strong> in five of the seven failures, complications were first detected by analysis of pH, pO₂ and pCO₂ before clinical symptoms became evident. Specificity for each of these parameters in drainage fluid samples was &gt; 94% in the presence of infection</td>
</tr>
<tr>
<td>Wheeler and Gilbert, 1999, UK</td>
<td>Surgery: colorectal resection (n = 102)</td>
<td>Investigation: intraoperative controlled water testing of anastomoses (maximum pressure exerted 30 cmH₂O)</td>
</tr>
<tr>
<td></td>
<td><strong>Aim:</strong> to examine whether controlled intraoperative water testing of left-sided colorectal anastomoses could determine in which patients a defunctioning ileostomy could be safely omitted</td>
<td><strong>Gold standard:</strong> routine water-soluble contrast enema on day 8 postoperatively. Clinical features not described</td>
</tr>
<tr>
<td></td>
<td></td>
<td><strong>Results:</strong> 21 (21%) patients failed the initial test; two of these developed clinical leak. Three (3%) patients failed a second test. Authors concluded a defunctioning ileostomy was avoided in 98% of patients and in 25/26 (96%) of low anastomoses; intraoperative testing is helpful but does not guarantee that an anastomosis will remain intact</td>
</tr>
<tr>
<td>Yamaguchi et al., 1998, Japan</td>
<td>Surgery: pancreatojejunostomy (n = 10)</td>
<td>Investigation: red litmus paper was used to detect transected pancreatic ductules on the cut surface of the pancreas</td>
</tr>
<tr>
<td></td>
<td><strong>Aim:</strong> a small prospective observational study (n = 10) in patients undergoing pancreatosenterostomy to evaluate a new simple technique for identifying a transected pancreatic ductule on the cut surface of the pancreas</td>
<td><strong>Results:</strong> the authors reported that no major postoperative anastomotic leak occurred</td>
</tr>
</tbody>
</table>
Methodology Group

Members

Methodology Programme
Director
Professor Richard Lilford
Director of Research
and Development
NHS Executive – West
Midlands, Birmingham

Chair
Professor Martin Buxton
Director, Health Economics
Research Group
Brunel University, Uxbridge

Professor Douglas Altman
Professor of Statistics
in Medicine
University of Oxford

Dr David Armstrong
Reader in Sociology as
Applied to Medicine
King’s College, London

Professor Nicholas Black
Professor of
Health Services Research
London School of Hygiene
& Tropical Medicine

Professor Ann Bowling
Professor of
Health Services Research
University College London
Medical School

Professor David Chadwick
Professor of Neurology
The Walton Centre for
Neurology & Neurosurgery
Liverpool

Dr Mike Clarke
Associate Director (Research)
UK Cochrane Centre, Oxford

Professor Paul Dieppe
Director, MRC Health Services
Research Centre
University of Bristol

Professor Michael Drummond
Director, Centre for
Health Economics
University of York

Dr Vikki Entwistle
Senior Research Fellow,
Health Services Research Unit
University of Aberdeen

Professor Peter Egan
Professor of Health Services
Research
University of York

Dr Nick Payne
Consultant Senior Lecturer
in Public Health Medicine
ScHARR
University of Sheffield

Professor Ewan B Ferlie
Professor of Public Services
Management
Imperial College, London

Professor Ray Fitzpatrick
Professor of Public Health
& Primary Care
University of Oxford

Dr Naomi Fulop
Deputy Director,
Service Delivery &
Organisation Programme
London School of Hygiene
& Tropical Medicine

Mrs Jenny Griffin
Head, Policy Research
Programme
Department of Health
London

Professor John Henderson
Economic Advisor
Department of Health, London

Professor Jeremy Grimshaw
Programme Director
Health Services Research Unit
University of Aberdeen

Professor Stephen Harrison
Professor of Social Policy
University of Manchester

Professor Theresa Marteau
Director, Psychology &
Genetics Research Group
Guy’s, King’s & St Thomas’s
School of Medicine, London

Dr Henry McQuay
Clinical Reader in Pain Relief
University of Oxford

Dr Nick Payne
Consultant Senior Lecturer
in Public Health Medicine
ScHARR
University of Sheffield

Professor Joy Townsend
Director, Centre for Research
in Primary & Community Care
University of Hertfordshire

Professor Kent Woods
Director, NHS HTA
Programme, & Professor
of Therapeutics
University of Leicester

Current and past membership details of all HTA ‘committees’ are available from the HTA website (see inside front cover for details)
# HTA Commissioning Board

## Members

| Programme Director | Professor Kent Woods  
| Personal/Institution | Director, NHS HTA  
| Personal/Institution | Programme, &  
| Personal/Institution | Professor of Therapeutics  
| Personal/Institution | University of Leicester  
| Chair | Professor Shah Ebrahim  
| Personal/Institution | Professor of Epidemiology  
| Personal/Institution | of Ageing  
| Personal/Institution | University of Bristol  
| Deputy Chair | Professor Jon Nicholl  
| Personal/Institution | Director, Medical Care  
| Personal/Institution | Research Unit  
| Personal/Institution | University of Sheffield  
|  | Professor Douglas Altman  
| Personal/Institution | Director, IC RF Medical  
| Personal/Institution | Statistics Group  
| Personal/Institution | University of Oxford  
|  | Professor John Bond  
| Personal/Institution | Director, Centre for Health  
| Personal/Institution | Services Research  
| Personal/Institution | University of Newcastle-upon-Tyne  
|  | Ms Christine Clark  
| Personal/Institution | Freelance Medical Writer  
| Personal/Institution | Bury, Lancs  
|  | Professor Martin Eccles  
| Personal/Institution | Professor of  
| Personal/Institution | Clinical Effectiveness  
| Personal/Institution | University of Newcastle-upon-Tyne  
|  | Dr Andrew Farmer  
| Personal/Institution | General Practitioner &  
| Personal/Institution | NHS R&D  
| Personal/Institution | Clinical Scientist  
| Personal/Institution | Institute of Health Sciences  
| Personal/Institution | University of Oxford  
|  | Professor Adrian Grant  
| Personal/Institution | Director, Health Services  
| Personal/Institution | Research Unit  
| Personal/Institution | University of Aberdeen  
|  | Dr Alastair Gray  
| Personal/Institution | Director, Health Economics  
| Personal/Institution | Research Centre  
| Personal/Institution | Institute of Health Sciences  
| Personal/Institution | University of Oxford  
|  | Professor Mark Haggard  
| Personal/Institution | Director, MRC Institute  
| Personal/Institution | of Hearing Research  
| Personal/Institution | University of Nottingham  
|  | Professor Jenny Hewison  
| Personal/Institution | Senior Lecturer  
| Personal/Institution | School of Psychology  
| Personal/Institution | University of Leeds  
|  | Professor Alison Kitson  
| Personal/Institution | Director, Royal College of  
| Personal/Institution | Nursing Institute, London  
|  | Dr Donna Lamping  
| Personal/Institution | Head, Health Services  
| Personal/Institution | Research Unit  
| Personal/Institution | London School of Hygiene &  
| Personal/Institution | Tropical Medicine  
|  | Professor David Neal  
| Personal/Institution | Professor of Surgery  
| Personal/Institution | University of Newcastle-upon-Tyne  
|  | Professor Gillian Parker  
| Personal/Institution | Nuffield Professor of  
| Personal/Institution | Community Care  
| Personal/Institution | University of Leicester  
|  | Dr Tim Peters  
| Personal/Institution | Reader in Medical Statistics  
| Personal/Institution | University of Bristol  
|  | Professor Martin Severs  
| Personal/Institution | Professor in Elderly  
| Personal/Institution | Health Care  
| Personal/Institution | University of Portsmouth  
|  | Dr Sarah Stewart-Brown  
| Personal/Institution | Director, Health Services  
| Personal/Institution | Research Unit  
| Personal/Institution | University of Oxford  
|  | Professor Ala Szczepura  
| Personal/Institution | Director, Centre for Health  
| Personal/Institution | Services Studies  
| Personal/Institution | University of Warwick  
|  | Dr Gillian Vivian  
| Personal/Institution | Consultant in Nuclear  
| Personal/Institution | Medicine & Radiology  
| Personal/Institution | Royal Cornwall Hospitals Trust  
| Personal/Institution | Truro  
|  | Professor Graham Watt  
| Personal/Institution | Department of  
| Personal/Institution | General Practice  
| Personal/Institution | University of Glasgow  
|  | Dr Jeremy Wyatt  
| Personal/Institution | Senior Fellow  
| Personal/Institution | Health Knowledge  
| Personal/Institution | Management Centre  
| Personal/Institution | University College London  

Current and past membership details of all HTA ’committees’ are available from the HTA website (see inside front cover for details)
Feedback

The HTA Programme and the authors would like to know your views about this report.

The Correspondence Page on the HTA website (http://www.ncchta.org) is a convenient way to publish your comments. If you prefer, you can send your comments to the address below, telling us whether you would like us to transfer them to the website.

We look forward to hearing from you.