Developing evidence-based clinical indicators: a state of the art methods primer

JAN MAINZ

The National Indicator Project and University of Aarhus, Denmark

Abstract

Objective. To describe steps in developing and testing clinical indicators based on state of the art methods in previous literature and experience in the Danish National Indicator Project.

Analysis. The development process includes a planning phase, where the clinical area to be evaluated is chosen and the measurement team selected and organized. The planning phase is followed by a development phase where clinical indicators are prioritized and selected by the measurement team on the basis of documentation and knowledge from the scientific literature. When clinical indicators have been selected, specific measure specifications should be designed, including inclusion and exclusion criteria for the target population, description of a risk adjustment strategy, identification of data sources, description of data collection procedures, and an analytical plan for data analyses. Before clinical indicators are implemented they should be tested for reliability and validity. Preliminary tests may identify areas requiring further modifications and specifications of the indicators.

Conclusion. Using clinical indicators for quality assessment represents an important approach to documenting the quality of care. Consumers of indicator information (clinicians, administrators, purchasers, regulators, and patients) need reliable and valid information for benchmarking, making judgments, and determining priorities, accountability and quality improvement. This underlines the fact that clinical indicators must be developed and tested with scientific rigor in a transparent process.

Keywords: clinical indicators, outcome measures, performance measures, quality improvement, quality of care

In most health care systems, a consensus is emerging that there is a need for quality measures. Various audiences may wish to use them to document the quality of care, make comparisons (benchmarking), make judgments and determine priorities, support accountability, support quality improvement, and provide transparency in health care [1,2].

Using clinical indicators for performance and outcome measurement is one way of measuring and monitoring the quality of care and services. In a companion paper in this journal, clinical indicators have been defined and characterized [3].

It is imperative that clinical indicators are meaningful, scientifically sound, generalizable, and interpretable. To achieve this, clinical indicators must be developed, tested, and implemented with scientific rigor [4].

This paper focuses on the development and testing of clinical indicators.

Steps in developing and testing indicators

The different steps required to develop and test clinical indicators are summarized in Table 1.

Each of these steps is discussed below. Specific examples of how they apply to clinical indicators for different diseases are provided to make the concepts more concrete, to illustrate some of the choices faced by developers of clinical indicators, and to facilitate understanding of the process of developing clinical indicators.

Choosing the clinical area to evaluate

There are many potential users of clinical indicators (clinicians, administrators, purchasers, regulators, and patients). Each user must clarify the purpose of the quality measurement effort. Several purposes may motivate quality measurement, such as regulation, purchasing, or quality improvement. Administrative quality improvement may arise from an organization's mission, values, and commitment, in response to patient complaints, or in response to payer or regulatory demands. The purposes will dictate the focus on particular clinical areas. Prioritizing among clinical areas for assessment may be based on various criteria, including the importance of the health care problem or disease and the opportunity for clinical interventions [2,4–6].

Table 1 Steps in the development and testing of clinical indicators

Planning phase	1. Choose the clinical area to evaluate
	Establish importance (high volume, cost, variation)
	Identify opportunities for clinical intervention
	2. Organize the measurement team
	Select group participants
	Organize and divide tasks
Development phase	3. Provide an overview of existing evidence and practice
	Present documentation and knowledge from the scientific literature for potential indicators
	Establish consensus about existing knowledge and practice
	4. Select clinical indicators and standards
	Select process indicators
	Select outcome indicators
	Identify prognostic factors (risk adjustment)
	Establish consensus and rating procedures
	5. Design measure specification
	Define indicators and standards
	Identify target population
	Determine inclusion and exclusion criteria
	Devise risk adjustment strategy
	Identify data sources
	Describe data collection procedures
	Develop a plan
	6. Perform pilot testing

The importance of the health care problem. McGlynn and others have suggested that a health care problem or a disease is important if it has a high volume, and is associated with high morbidity and mortality and is costly to treat.

Local or national epidemiological data can be used to determine the prevalence of disease in a population. Mortality rates, significant use of health services, and costly treatment are other criteria that have been cited to support a focus on a particular condition for quality measurement [2,4–6]. Disease-specific mortality rates for a variety of conditions are available in most countries [1,5,6]. For example, in Denmark, such data are available from the Central Patient Registry [6]. Health services utilization rates for a particular condition and costs of treatment have also been cited by McGlynn and others as important indicators of importance and, when available, may be used as additional criteria for choosing conditions for quality measurement [1,5,6].

The higher the prevalence and incidence of a condition or frequencies of procedures or outcomes, the more likely it is that it will be possible to identify an adequate number of cases for quality measurement.

Opportunities for clinical interventions. Another important criterion for choosing the clinical area for review is the opportunity for interventions related to the health care problem or disease. Clinical indicators are most useful if the processes and outcomes being assessed can be influenced by clinical interventions in terms of quality improvements efforts. For each health problem or disease it is therefore important to consider what actions are available to improve the quality of care. Conditions with high volumes might be chosen for indicator monitoring. Rare conditions have often not had many clinical trials for logistical reasons, and do not have evidence-based guidelines developed. Therefore, it is difficult to know which processes might be improved by monitoring. In the Danish National Indicator Project, stroke and lung cancer were selected, partly because evidence-based clinical guidelines were available. Greater priority should be given to clinical areas where there is evidence that the quality of care is either variable or substandard, so that areas with a substantial potential for quality improvement are chosen. In the Danish National Indicator Project, lung cancer was included because mortality rates indicated that Denmark had a high mortality rate compared with other countries [6].

Organize the measurement team

Select team members. The measurement team may be stronger if it represents different perspectives. Clinicians can ensure that appropriate clinical indicators are selected together with standards to evaluate whether desirable performance or outcome rates are obtained. Often it may be relevant to select a multidisciplinary team of clinicians, including, for example, doctors from different specialties, nurses, physiotherapists, and occupational therapists.

Clinicians being evaluated will have more confidence in the indicators developed if measurement team members include clinicians who are widely recognized and respected. Relevant credentials include appropriate professional training and an active role in professional societies.

For instance, when selecting clinicians for a multidisciplinary team to develop clinical indicators for lung cancer, it might be relevant to include thoracic surgeons, internists specializing in pulmonary diseases, oncologists, nurses, physiotherapists, and psychologists [6]. When clinical indicators for schizophrenia are developed, psychiatrists, nurses specializing in psychiatry, psychologists, and social workers might be relevant to include in the measurement team [6].

If the quality measurement is conducted in a nation, a region, state, or county, it might be best to draw the team members from different geographical areas, from urban and rural locales, and from different types and sizes of organization.

The size of the team, however, should be kept small, so that discussions and teamwork do not become unwieldy. In the Danish National Indicator Project, the teams selected consist of eight to 14 professional clinicians appointed by different scientific societies [6].

Quality of care researchers together with clinical epidemiological expertise can help to ensure methodological integrity of the clinical indicators and a sound approach to data collection and data analysis. Like multi-center clinical trials, multi-organizational measurement involves methodological issues that require highly trained scientists.

According to the selected clinical area, it might be relevant to include patient or administrator representatives. Patients and administrator representatives can contribute to the selection and prioritization of some clinical indicators. For complex medical technical indicators, their contribution will be less. The Clinical Standards Board of Scotland usually has patients and organizational representatives 'on board' their selecting teams [7].

Organize teamwork. Once the measurement team has been selected, the work in the group can be organized, including planning relevant meetings, the creation of smaller working subgroups, and the delegation of tasks to individual group participants. In the Danish National Indicator Project we found it helpful that all meetings in the measurement team were followed by exhaustive summaries of decisions made by the group.

In nationwide measurement projects, reimbursement of clinical departments for clinicians, time used in meetings, and other tasks may facilitate the work within the measurement team [6].

Provide an overview of existing evidence and practice

Clinical indicators should be based on scientific evidence or consensus among health professionals. Preferably, clinical indicators should be based on research evidence rather than on expert opinions or clinical experience alone. The level of evidence supporting each clinical indicator is transparent when the strength of that evidence is described before the indicators are selected.

Providing an overview of existing knowledge from the scientific literature and practice allows the measurement team to take into account the strength of evidence when choosing clinical indicators. There are numerous systems for rating the strength of scientific evidence [8,9]. The highest level of evidence is obtained by meta-analysis of randomized controlled trials and evidence from at least one randomized controlled trial ('A'-evidence). 'B'-evidence is obtained for controlled studies without randomization or quasi-experimental studies. 'C'-evidence relates to different epidemiological studies such as case-control studies. Finally, 'D'-evidence refers to evidence based on different expert opinions. Using a rating scheme to summarize the strength of evidence enables the measurement team to describe the evidence of clinical indicators [9].

Literature databases such as the Cochrane Collaboration or Medline, and the compendium 'Clinical Evidence' are important sources for determining the strength of evidence for clinical indicators (available online at http://www.cochrane.org, http://www.ncbi.nlm.nih.gov/PubMed, and http://www.clinicalevidence.org).

Select clinical indicators and standards

A clinical indicator is a measure that assesses a particular health care process or an outcome [3]. In other words a clinical indicator is a tool for producing a quantitative measurement of quality of care. However, simply knowing the level of an indicator does not reveal whether or not it is acceptable. A judgment of the acceptability of a performance or an outcome rate must be made in relation to the purpose for which it should be used. The final stage in measuring health care quality is applying a standard of quality that embodies acceptability of a particular performance or outcome rate. If a desired attribute of care falls below the standard or an undesired attribute of care rises above this level, further evaluation or action is triggered. There may be instances where standards are not established and must be based on a preliminary data collection when benchmarks are available. Before the establishment of standards, providers or populations can be compared to determine relative quality of care, but adherence to a standard cannot be determined. The strength of evidence for both the clinical indicators and related standards of care can be described. Multiple indicators can be used to evaluate the quality of care for most conditions. It is best to select clinical indicators supported by evidence indicating that they reflect the process and outcome being evaluated. If no scientific evidence is available, clinical indicators can be selected on the basis of consensus among health professionals.

Clinical process indicators. Measuring the quality of a process of care requires determination of whether clinicians are adhering to practices to achieve the best outcomes for patients. The linkage of a process to outcomes must ideally have been demonstrated scientifically.

To determine whether to measure a specific process of care it is helpful to review the strength of scientific literature supporting the fact that inclusion of this process in the process measure will affect outcomes. The processes of care measured should be those demonstrated to cause a higher probability of achieving a desired outcome. The strength of the evidence for an indicator will determine its scientific soundness or the likelihood that improvements in the clinical indicator will produce consistent and credible improvement in quality of care. Clinical

Table 2 Examples of clinical process indicators

Indicator concept	Definition of indicator	Type	Standard of care	Time of intervention/ treatment	Level of evidence
Organization of treatment (stroke)	Proportion of patients treated/ rehabilitated in stroke units	Process	More than 90% of patients with acute stroke should be treated and rehabilitated in a stroke unit	<24 hours after admission	A
Treatment (lung cancer)	Proportion of patients treated	Process	Treatment rate ≥70%	Discharge	B
Resection (lung cancer)	Proportion of patients resected	Process	Resection rate ≥25%	Discharge	B

A, evidence from at least one randomized controlled trial; B, evidence from quasi-experimental or non-randomized controlled studies.

process indicators based on A-evidence are most credible, but there might be arguments for selecting indicators with a lower strength of evidence.

Initially, when evidence links a process to better outcomes it may appear that the standard for a proportion of patients so treated should be 100%. However, there are reasons why this is not always the case, depending on how well the denominator of eligible patients can be defined.

Table 2 illustrates an example of a clinical process indicator for stroke. The indicator measures the proportion of patients treated and rehabilitated in a stroke unit [6]. The evidence supporting this indicator is strong (A-evidence), since meta-analyses of randomized controlled trials have demonstrated the effects of stroke units on outcomes of care [10,11]. Compared with treatment in departments of internal medicine, treatment in stroke units was associated with lower mortality [odds Ratio (OR)=0.83; 95% confidence interval (CI) 0.71–0.97]. On the basis of this evidence it has been recommended that all patients with acute stroke be treated at specialized stroke units [10,11]. Table 2 illustrates, however, that the selected standard suggests that >90% of patients with acute stroke should be treated at stroke units. The reason for this discrepancy is that a smaller subgroup of patients (~10%) with acute stroke will not benefit from treatment at specialized stroke units because they are deceased by the time of admittance to hospital or intensive care unit [6].

Table 2 also illustrates two clinical process indicators for lung cancer: proportion of patients who are actively treated and proportion of patients who are resected. The active treatment rate expresses the proportion of patients who were offered a concrete treatment (e.g. resection, chemotherapy, or radiation therapy).

The scientific literature indicates that active treatment for all patients with lung cancer is important with regard to survival and the patients' quality of life (B-evidence) [12–14]. According to the literature, >70% of patients with primary lung cancer should be offered an active treatment (C-evidence).

Cohort studies have shown that a high resection rate is associated with high survival (B-evidence) [12–14]. The studies indicate that >25% of patients should be resected.

Although the presented indicators for lung cancer are associated with B- and C-evidence, they might be regarded as important by a measurement team when evaluating the treatment for lung cancer, because it is important from a clinical perspective to evaluate this clinical practice even though no randomized controlled studies have been conducted [6].

For most process indicators, risk adjustment plays a smaller role than it does for outcome measurement. For some process measures, however, risk adjustment may reveal that patient factors are influencing a measure. The more closely an indicator measures the actual process of care delivered rather than patient adherence or other factors, the less risk adjustment will be needed [4].

Clinical outcome indicators. Multiple factors contribute to health care outcomes [3]. When evaluating outcome indicators, the adequacy of controls for differences in case mix and the adequacy of controls for other covariates are important criteria.

Case mix or severity-of-illness adjustment allow for a 'fair' comparison of health outcomes to ensure that any observed differences can be attributed to the health care interventions and not to differences between the populations included [5]. Patients who die or recover more slowly may not have received poorer quality care, but have been at higher risk for the outcomes before treatment.

Table 3 lists different outcome indicators. Intermediate outcome indicators reflect changes in biological status that affect subsequent health outcomes, and can be regarded as short-term outcomes [3]. For example, large RCTs in people with type 1 and type 2 diabetes have found that risk of development or progression of complications increases progressively as hemoglobin A1c (HbA1c) increases above the non-diabetic range (≥7.0 mmol/l) (A-evidence) [15]. HbA1c therefore reflects important health outcomes that can only be measured after years. It is considered desirable for most diabetics (≥90%) to have a HbA1c <7.0 mmol/l [15].

The last outcome indicator presented in Table 3 refers to 30-day mortality for stroke patients. Stroke is associated with high mortality. The literature suggests that case fatality should

Table 3 Examples of outcome indicators

Indicator concept	Indicator definition	Type	Standard of care	Time of intervention/ treatment	Level of evidence
Blood glucose control	Proportion of diabetics with HbA1c <7.0 mmol/l	Intermediate outcome	≥90% should have HbA1c <7.0 mmol/l	Every third month	A
Mortality (stroke)	30-day and 3-, 6- and 12-month mortality	Outcome	<20% should have a 30-day mortality	30 days after stroke	B

A, evidence from at least one randomized controlled trial; B, evidence from quasi-experimental or non-randomized controlled studies.

be <20%. Prognostic factors according to the literature are age, sex, previous stroke, diabetes, and social status [16].

The selection process. The measurement team selects clinical indicators in a consensus process. For a specific clinical area, team members can suggest multiple clinical indicators. The team may use several methods for selecting among them. Criteria may include the indicator's importance, strength of evidence for the indicator and the indicator's validity and reliability, and the flexibility of obtaining the indicator data. The team may apply various processes to make decisions based on these criteria, including consensus, majority vote, or prioritization according to average ratings derived using formal techniques such as those used by RAND for appropriateness criteria [17]. The Danish National Indicator Project has chosen to select indicators by team consensus in order to enhance the credibility of indicators among intended users [6]. Each indicator should be considered by the measurement team according to importance, strength of evidence, validity and reliability, and feasibility. Within the selection of indicators, different rating procedures are available [17].

Design measure specifications

When potential indicators have been selected, the next step is to design a reliable and valid measure that can be implemented consistently. It is helpful when measure specifications are described in specific manuals in order to minimize inter-rater variation [6].

Define the clinical indicator and standard. The first step is to describe each indicator in detail. Examples of indicators are described in Tables 2 and 3. Some measures are dichotomous where the answer is either 'yes' or 'no', such as whether a patient with lung cancer has undergone resection. Dichotomous measures, when aggregated for the relevant population to which a quality indicator is applied, are expressed as proportions, with a given numerator and denominator, such as the proportion of patients with lung cancer who have had resection (Table 2). Dichotomous measures can also be used to generate rates, which are proportions within a given time period. An example of a rate is the proportion of stroke patients treated in a stroke unit within 24 hours (Table 2) [18].

Other measures are continuous, such as the number of minutes from arrival at the health care facility until a myocardial infarct patient received thrombolytic treatment. Indicators based on these measures would be expressed as means for the relevant population. All details of the indicator should be specified. For instance, for the above-mentioned indicator, the concept 'stroke unit' should be defined. A stroke unit can be described as a hospital unit or part of a hospital unit that solely, or almost solely, treats or rehabilitates patients with stroke [6].

When the indicators have been defined, standards describing what are acceptable performances or outcomes should be specified. As described above, standards can be derived from the academic literature or from consensus within the measurement team. The combination of evidence from the scientific literature and clinical experience may be helpful for the measurement team to derive relevant and realistic standards.

Identify target population. The target population refers to the patient group whose care the clinical indicator is designed to assess. Specific inclusion and exclusion criteria have to be defined. It should also be decided whether the selection should be based on confirmed diagnoses, or symptoms or signs. Whether prevalent or incident cases (or both) are included should also be taken into consideration. It might be relevant to describe upper or lower age limits. Finally, decisions about the time period for measurement should be taken.

Determine risk adjustment strategy. As described above, prognostic factors have to be identified and described together with clinical outcome indicators. The data required for risk adjustment are indicated above. Specific data definitions should be described for all included prognostic factors (patient characteristics and comorbidities).

Identify data sources. Once the clinical indicators have been defined, the prognostic factors for risk adjustments derived, and the target population identified, the measurement team must state how the data should be obtained.

Different data sources are used for quality measurement. Administrative data are generated to support reimbursement activities. This data source is readily available and inexpensive to collect, but it might be unreliable because of coding problems, and therefore lacks specificity and detail. Clinical data include medical record data and data from clinical research databases. Medical record data are the most complete sources of information on diagnosis, treatment, and clinical outcomes. They are, however, expensive to obtain. Data from clinical research

databases are often valid and reliable, if such databases are available in the chosen clinical area. Primary data, in terms of prospectively collected clinical data that are collected for a particular quality measurement purpose, are the most specific and can define exactly what data are required. Primary data can also include survey data, from patients to access attitudes, behavior, knowledge, and outcome. Since primary data represent data that are not readily available, such data are expensive to collect, but are often a valid and reliable information source.

In quality measurement, data collection can be made part of routine care by standardizing the documentation of patient characteristics and care delivery that clinicians and administrators are already recording while delivering care and services. Rubin *et al.* have noted that this is particularly useful for the efficient use of process indicators [2,4], and may reduce missing data and the additional cost of data collection. When electronic medical records become widespread, standard data specifications useful for quality assessment can be incorporated into these systems. This would eliminate duplicative clinical data collection for the purposes of clinical care and quality assessment [2,4].

Describe data collection procedures. Description of detailed specifications for data collection allows different health institutions (e.g. hospitals) to implement quality measures in a consistent way so that results may be fairly compared.

A protocol for scoring the measures can be developed according to the measure specifications. A plan for handling missing data or data outside of a logical range must also be included in the specifications.

Develop an analytical plan. As part of the development process, Rubin *et al.* recommend developing a detailed plan for how the measures are to be analyzed and how statistical and clinical significance should be determined [2,4]. The analytical plan could include a detailed description of the population, an assessment of the distribution of the data, how missing data are handled, and the statistical analyses of tests to be used. The measurement team may wish to consider what are clinically significant differences among groups rather than simply statistically significant differences. Comparing the clinical significance of differences among groups helps answer a question that data users frequently ask: 'what is good quality of care?'.

Preliminary testing

Before clinical indicators are implemented they should be thoroughly tested. Preliminary tests may identify areas that require further specifications of the quality measures.

Reliability of a clinical indicator expresses the extent to which repeated measurements of a stable phenomenon by different providers and instruments, at different times and places, obtain similar results. Reliability is important for comparing groups or comparing the same group over time periods. Reliability can be tested as inter-rater reliability, where different people or methods provide data on the same indicator. Reliability can also be tested as internal consistency, for which two indicators, expected to measure the same

aspect of quality of care, are compared. Measuring inter-rater reliability, internal consistency, and test–re-test reliability allows users to determine if the data collection methods are precise enough to provide reproducible results. These methods assess data quality powerfully and identify whether the measure and data collection procedures are well specified.

Validity determines the degree to which an indicator measures what it is intended to measure, that is whether the results of a measurement corresponds to the true state of the phenomenon being measured [2–4]. Validity can be tested by confirming that the scores of a measure are linked to specific outcomes, and that the measure can reflect good and bad quality.

Conclusion

This paper has outlined the steps in a state-of the-art development process of clinical indicators. Performance and outcome indicators represent the only way to obtain quantitative data on the quality of care for quality improvement.

Clinical areas should be selected according to the importance of health care problems and known opportunities for clinical interventions. In the development process, representatives that can contribute significantly should be involved. Because clinical indicators measure the cornerstones of clinical practice, recognized and respected clinicians play a central role. Nation-wide or regional quality measurement efforts may wish to include team members from different geographical areas for increased acceptance and credibility.

A development process based on scientific principles will create more valid indicators and will also increase acceptance and credibility of the resulting data. For each selected indicator and related standard of care, transparency about measures, whether evidence or consensus based, and the level of evidence available, will assist users in assessing how much the resolving data matter to them.

For some process measures and for all outcome measures, prognostic factors (patient characteristics and comorbidities) are generally needed to define a risk adjustment strategy.

As part of the development process, exhaustive and exclusive measure specifications should be described, including specific definitions of the clinical indicators and standards, identification of the target population and data sources; and development of an analytical plan with description of statistical and clinical significance of the results will be assessed when comparing groups or comparing a group to a standard.

Before implementation of clinical indicators, preliminary tests for reliability and validity can provide users with information about the trustworthiness and usefulness of data.

Indicator measurement can be made most efficient when incorporated into routine patient care as part of clinicians' and administrators' documentation of required information on patient characteristics and care delivery, which is already recorded for clinical purposes.

When electronic medical records become widespread, data specification for quality assessment can be incorporated into such systems, which will avoid duplicative clinical data collection and reduce additional cost.

References

1. US Department of Health and Human Services, Public Health Service, and Agency for Health Care Policy and Research. AHCPR Publication No. 95-0046, March 1995.

2. Rubin HR, Pronovost P, Diette GB. The advantages and disadvantages of process-based measures of health care quality. *Int J Qual Health Care* 2001; **13**: 469–474.

3. Mainz J. Defining clinical indicators for quality improvement. *Int J Qual Health Care* 2003, 523–530.

4. Rubin HR, Pronovost P, Diette GB. From a process of care to a measure: the development and testing of a quality indicator. *Int J Qual Health Care* 2001; **13**: 489–496.

5. McGlynn EA, Asch SM. Developing a clinical performance measure. *Am J Prov Med* 1998; **14**: 14–21.

6. Mainz J, Bartels PD, Laustsen S *et al*. The National Indicator Project to monitoring and improving of the medical technical care. *Ugeskr Laeger* 2001; **163**: 6401–6406.

7. Clinical Standards Board of Scotland. *Colorectal Cancer Services*. Edinburgh: CSBS, 2002.

8. West S, King V, Cary TS *et al*. Systems to *Rate the Strength of Scientific Evidence*. Evidence Report/Technology Assessment No. 47. AHRQ Publication No. 02-E016. Rockville, MD: Agency for Healthcare Research and Quality, April 2002.

9. Eccles M, Fremantle N, Mason J. North of England evidence-based guidelines development project: methods of developing guidelines for efficient drug use in primary care. *Br Med J* 1998; **316**: 1232–1235.

10. Langhorne P, Dennis M, eds. *Stroke Units: An Evidence-Based Approach*. London: BMJ Books, 1998.

11. Stroke Unit Trialists' Collaboration. *Organised Inpatient (Stroke Unit) Care for Stroke (Cochrane Review)*. In The Cochrane Library, Issue 1. Oxford: Update Software, 2001.

12. Breathnach OS, Freidlin B, Conley B *et al*. Twenty-two years of phase III trials for patients with advanced non-small-cell lung cancer: sobering results. *J Clin Oncol* 2001; **19**: 1734–1742.

13. Engeland A, Bjorge T, Haldorsen T, Tretli S. Prognosis of patients with lung cancer diagnosed in Norway, 1954–93. *Cancer Causes Control* 1998; **9**: 57–65.

14. Laroche C, Wells F, Coulden R *et al*. Improving surgical resection rate in lung cancer. *Thorax* 1998; **53**: 445–449.

15. The Diabetes Control and Complications Trial Research Group. The effect of intensive treatment of diabetes on the development and progression of long-term complications in insulin-dependent diabetes. *N Engl J Med* 1993; **329**: 977–986.

16. Aboderin I, Venebles G. Pan European consensus meeting on stroke management. *J Intern Med* 1996; **240**: 173–180.

17. Brook RH, McGlynn EA, Cleary PD. Measuring quality of care. *N Engl J Med* 1996; **335**: 960–970.

18. Sackett D, Haynes RB, Guyatt GH, Tugwell P. *Clinical Epidemiology. A Basic Science for Clinical Medicine*. Boston, MA: Little, Brown and Company, 1991.