How can hospital performance be measured and monitored?

August 2003
ABSTRACT

Health Evidence Network (HEN) synthesis report on hospital performance

Measurement is central to the concept of hospital quality improvement; it provides a means to define what hospitals actually do, and to compare that with the original targets in order to identify opportunities for improvement. The principal methods of measuring hospital performance are regulatory inspection, public satisfaction surveys, third-party assessment, and statistical indicators, most of which have never been tested rigorously.

This report is HEN’s response to a question from a decision-maker. It provides a synthesis of the best available evidence, including a summary of the main findings and policy options related to the issue.

HEN, initiated and coordinated by the WHO Regional Office for Europe, is an information service for public health and health care decision-makers in the WHO European Region. Other interested parties might also benefit from HEN.

This HEN evidence report is a commissioned work and the contents are the responsibility of the authors. They do not necessarily reflect the official policies of WHO/Europe. The reports were subjected to international review, managed by the HEN team.

When referencing this report, please use the following attribution:

Keywords
QUALITY INDICATORS, HEALTH CARE - STANDARDS
HOSPITALS – STATISTICS – STANDARDS
QUALITY OF HEALTH CARE
DELIVERY OF HEALTH CARE
DATA COLLECTION – METHODS
DECISION SUPPORT TECHNIQUES
HEALTH POLICY
EUROPE

© World Health Organization 2003

All rights reserved. The Regional Office for Europe of the World Health Organization welcomes requests for permission to reproduce or translate its publications, in part or in full.

The designations employed and the presentation of the material in this publication do not imply the expression of any opinion whatsoever on the part of the World Health Organization concerning the legal status of any country, territory, city or area or of its authorities, or concerning the delimitation of its frontiers or boundaries. Where the designation “country or area” appears in the headings of tables, it covers countries, territories, cities, or areas. Dotted lines on maps represent approximate border lines for which there may not yet be full agreement.

The mention of specific companies or of certain manufacturers’ products does not imply that they are endorsed or recommended by the World Health Organization in preference to others of a similar nature that are not mentioned. Errors and omissions excepted, the names of proprietary products are distinguished by initial capital letters.

The World Health Organization does not warrant that the information contained in this publication is complete and correct and shall not be liable for any damages incurred as a result of its use. The views expressed by authors or editors do not necessarily represent the decisions or the stated policy of the World Health Organization.
Summary

The issue
Measurement is central to the concept of quality improvement; it provides a means to define what hospitals actually do, and to compare that with the original targets in order to identify opportunities for improvement.

Findings
The principal methods of measuring hospital performance are regulatory inspection, public satisfaction surveys, third-party assessment, and statistical indicators, most of which have never been tested rigorously. Evidence of their relative effectiveness comes mostly from descriptive studies rather than from controlled trials. The effectiveness of measurement strategies depends on many variables including their purpose, the national culture, how they are applied and how the results are used.

Inspection of hospitals measures minimal requirements for the safety of patients and personnel; it does not foster innovation or information for consumers or providers.

Surveys usually address what is valued by patients and the general public. Standardized surveys measure specific domains of patient experience and satisfaction. There are also standardized surveys that reliably measure hospital performance against explicit standards at a national level.

Third party assessments may include measurement by standards, by peer review or by accreditation programmes. ISO standards assess compliance with international standards for quality systems, rather than hospital functions per se. Peer review is generally supported by clinical professions as a means of self-regulation and improvement, and does not aim to measure the overall performance of hospitals. Accreditation programmes are managed by independent agencies in several countries. They focus on what may be improved rather than on failures, and are oriented toward the patient, the clinical procedures, outcome and organizational performance. These programmes require substantial investments, and there is ample evidence that hospitals rapidly increase compliance with published standards and improve organizational processes in the months prior to external assessment. There is less evidence that this brings benefits in terms of clinical process and patient outcome.

Statistical indicators can suggest issues for performance management, quality improvement and further scrutiny; however, they need to be interpreted with caution. Much of the current evidence on the effectiveness of performance indicators is based on observational or experimental data. Some experience suggests that indicators such as guidelines to standardize management of common conditions may reduce length of stay and episode costs without detriment to clinical outcome. The publication of performance statistics as “league tables” aims to encourage improvement, to empower patient choice and to demonstrate a commitment to transparency. Evidence suggests that this increases public interest and management attention to data quality, but it does not appear to have much effect on performance.

Policy considerations
Systems for measuring hospital performance should be published in a national or regional plan for quality and performance management, and clarify the roles and values of stakeholders.

The design of performance measurement systems should aim to improve hospital performance, rather than to identify individual failures. Systems should not rely on single sources of data but should use a range of information. Consumers should be prominently involved, and the results of assessments should be transparent and freely accessible to the public.
What is performance and where is it measured?

“Performance” must be defined in relation to explicit goals reflecting the values of various stakeholders (such as patients, professions, insurers, regulators). In reality, however, very few performance measurement systems focus on health outcomes valued by customers. “Measurement” implies objective assessment but does not itself include judgement of values or quality; these may be added by those who later present and interpret the data.

At the system level, improvement in such areas as health priority setting, system planning, financing and resource allocation, professional recognition and overall quality management often become important aims of health reforms. At the national level, many countries, such as Ireland (1), Denmark (2), the United Kingdom (3), and the Germany (4), have developed frameworks for performance assessment and improvement.

At the European level much work has been done to summarize data on hospital performance and quality assurance policies in the European Union (5), accession states (6) and other WHO Member States. General recommendations on the development and implementation of quality improvement systems in health care were made to health ministers by the Council of Europe in 1997 (7), and best practices in the efficient and effective delivery of services were published by the European Commission in 1999 (8).

At the global level, findings concerning health systems performance measurement in 192 Member States were summarized in the WHO World Health Report 2000 (9). This document sets out a framework for evaluating and improving performance of health systems in four key functions: providing services, creating resources, financing and oversight.

Hospital performance may be defined according to the achievement of specified targets, either clinical or administrative (10). Ultimately, the goal of health care is better health, but there are many intermediate measures of both process and outcome. Targets may relate to traditional hospital functions, such as diagnosis, treatment, care and rehabilitation as well as to teaching and research. However, both the definition and the functions of hospitals are changing, as emphasis shifts from inpatient care to ambulatory care, community outreach programmes and health care networks (11). Hospital performance may thus be expected to include elements of community care and public health, as well as social and employment functions. These dimensions of hospital performance have been analysed in the European context (12).

Measurement is central to the concept of quality improvement; it provides a means to define what hospitals actually do, and to compare that with the original targets or expectations in order to identify opportunities for improvement. Hospitals have many targets and many stakeholders; these may be seen as clusters of values and aims behind performance measurement (13), in such areas as:

- **Research**: Data about structure, activities and effectiveness can be used to study the link between organization and performance, and to inform planning and system development.
• **Service improvement:** Purchasers and providers can compare performance within and among hospitals to stimulate and measure change.

• **Referrer and patient choice:** Patients and their referrers can use information such as waiting times, outcomes and patient experiences in choosing a provider.

• **Resource management:** Purchasers and provider managers need data on performance, costs and volume of activity in order to decide on the best use of resources.

• **Accountability:** Politicians and the public increasingly demand transparency, protection and accountability for performance.

Hospitals need positive incentives to provide timely, accurate and complete data to external assessment programmes. If such programmes are perceived to have intrinsic value to the organization (for example, in staff motivation, team building; clinical and professional development or risk management), hospitals have less need for financial or market incentives to participate. Conversely, neither individuals nor hospitals are keen to provide information which might lead to public blame, litigation, and loss of staff, authority and trade. Many performance measurement systems assume a common culture of transparency, professionalism and accountability that motivates cooperation.

**Sources for this review**

Unlike clinical practice and health technology, performance measurement (along with many health policy initiatives) has rarely been submitted to rigorous scientifically controlled trials. The best available evidence tends to be empirical successes and failures from other countries or regions. Hence this synthesis draws heavily on information published by health ministries (usually on intentions rather than achievements), and on independent research and evaluations by expert groups, particularly within the European Region.

**What are the methods of performance measurement?**

The methods used for performance measurement and quality improvement have not been rigorously evaluated within or across countries, largely because they are complex interventions which are not easily isolated and measured. The evidence to support these strategies is mostly based on descriptive studies or expert reports and on respected authority. There are in principle five different types of measurement of hospital performance, of which the first four will be dealt with here:

• regulatory inspection
• surveys of consumers’ experiences
• third-party assessments
• statistical indicators
• internal assessments.

**Inspection**

Most countries have statutory inspectorates to monitor compliance of hospitals with published licensing regulations. More specialized functions include fire, hygiene, radiation, medical devices and medicines, and some countries include infection control and blood transfusions. Inspections standards have legal authority and are transparent, but by the same token are not easily updated. Standards address the minimal legal requirements for a health care organization to operate and care for patients; they do not usually address clinical process or hospital performance. Licensing inspections often apply only to new hospitals, particularly in the private sector; where relicensing is applied, certificates may be issued on payment of a fee with minimal or no inspection. When assessment is managed locally by a governmental entity or its designated agent, there may be little national consistency or aggregation of reports, and when it is highly centralized, results are often not shared with staff or patients.
How can hospital performance be measured and monitored?
WHO Regional Office for Europe’s Health Evidence Network (HEN)
August 2003

Some governmental agencies — for example, the Joint Commission in the United States — define the standards for hospital licenses, but issue them on the basis of assessments made independently by accreditation programmes that they monitor for conformity. In the United Kingdom, The Commission for Health Improvement (CHI) was designed to inspect arrangements for “clinical governance” in public hospitals in England and Wales. CHI published no standards for self-assessment and formed no reciprocation with independent or private organizations. Their reports are detailed and public.

Inspection of hospitals induces conformity, and measures performance in terms of minimal requirements for safety. It does not foster innovation or information for consumers or providers.

Consumer surveys

Standardized surveys of patients and relatives can reliably measure hospital performance against explicit standards at a national level. Hospital performance is becoming more focused on health education, patient empowerment, comfort, complaint mechanisms and continuity of care. Some governments and intergovernmental organizations seek to make patients more aware of their rights – and to increase their sometimes very low expectations – by publishing patients charters and by legislating the protection of patients’ rights. Thus, consumer surveys assessing the experience of health care and outcomes as perceived by patients and their families carry added weight. Some countries (including France and the United Kingdom) and most accreditation programmes require institutions to make systematic assessments of their patients’ perceptions. Surveys range from local pencil-and-paper surveys outside a clinic to national stratified sample surveys. National surveys are often managed under contract by independent organizations using validated tools to obtain reliable data; published results may identify the performance of individual hospitals.

Advantages of this method are that it identifies what is valued by patients and the general public, and standardized surveys can be tailored to measure specific domains of experience and satisfaction. However, traditional satisfaction surveys have been methodologically weak, and focused on the agenda of clinicians and managers rather than patients. A review of 195 published studies suggested that few patient surveys were both valid and reliable (14), and governments may be reluctant to publish adverse results for public hospitals.

Many patients have low expectations and are too readily satisfied; systematic measurement of their experience is a more sensitive indicator of empowerment (15). Researchers at Harvard Medical School developed and tested a standardized instrument to measure patients’ concerns and experience. It was first used at a national level to interview hospital inpatients and relatives by telephone in the United States (16), and has since been used as the so-called Picker Questionnaire in Australia, Canada (17) and various European countries (18, 19, 20). Favourable Picker scores have shown correlations to significantly reduced complications and unexpected deaths in Michigan hospitals (21), and low scores were associated with lower health status among patients with acute myocardial infarction in New Hampshire (22).

In England, all hospitals are required to commission their own local surveys each year, including a standard set of questions for national performance monitoring and benchmarking. Results are submitted to the Department of Health for use in the National Performance Assessment Framework. Aggregated results are published on the Internet (23), and financial incentives are offered for demonstrably patient-centred care (24).

Studies in, for example, France (25), Greece (26), Poland (27), Sweden (28) and the United Kingdom (29) have shown that inter-hospital comparisons are feasible at a local or regional level.

Third party assessment

A research project funded by the European Union (30) identified systematic approaches linking national or international standards to local practices of private or public hospitals. These approaches have been compared in a number of studies of standards and methods used by industry-based (ISO, Baldrige) and
How can hospital performance be measured and monitored?

WHO Regional Office for Europe’s Health Evidence Network (HEN)
August 2003

health-care-based (peer review, accreditation) programmes (31, 32, 33, 34). The programmes, which are voluntary and independent to varying degrees, use explicit standards to combine internal self-assessment with external review by visits, surveys, assessments or audits (35). As the previously cited survey of 195 studies (14) says: “Considering the amount of time and money spent on organizational assessment, and the significance of the issue to governments, it is surprising that there is no research into the cost-effectiveness of these schemes.”

ISO Standards

International Organization for Standardization (36) certification measures hospital performance in terms of compliance with international standards for quality systems, rather than in terms of hospital functions and objectives. Details of assessments are not publicly available. ISO developed a series of standards (ISO 9000) originally for the manufacturing industry (medicines, medical devices) that have been used to assess quality systems in specific aspects of health services and hospitals and clinics. Hospitals (or, more commonly, parts of them) are assessed by independent auditors who are themselves regulated by a national “accreditation” agency. The theoretical advantage is that ISO certification is internationally recognized in many other service and manufacturing areas, but ISO 9000 standards relate more to administrative procedures rather than to hospital performance. Furthermore, the terminology of the standards is difficult to relate to health care, and interpretations vary among national agencies (37). The audit process tests compliance with standards and is not intended for organizational development. Few whole hospitals have been ISO certified and few countries have a national register of these hospitals.

The ISO 9000 standards for quality systems were adapted in 2000 to become more easily applied to health care and to include the assessment of outcomes and consumer satisfaction. There are initiatives in the United States (led by the major motor manufacturers who purchase health care for their employees) and in Europe (led by CEN) to interpret quality standards for health care. ISO15189 is becoming the international standard for medical laboratories and includes issues of clinical judgement, process and outcome.

Peer review

Peer review is a closed system for professional self-assessment and development. Reciprocal visiting is driven by professional (often single-discipline) organizations and has a long tradition as a form of peer review, especially for the recognition of training posts. It is endorsed by clinical professions as a means of self-regulation and clinical improvement, and is integrated with undergraduate, specialty and continuing professional development. Reciprocal visiting has also been applied to service development, such as in the hospital specialties programme in the Netherlands (38). Limitations of the method include its basis in specialties, as opposed to whole hospitals, and the confidentiality of its results.

Peer review schemes could provide a source of standards and assessments to harmonize professional and human resource management within and between countries with reciprocal recognition of training.

Accreditation

Accreditation programmes measure hospital performance in terms of compliance with published standards of organizational – and, increasingly, clinical – processes and results. They are mostly independent and aimed at organizational development more than regulation but could contribute reliable data to national performance measurement systems. They are independent, voluntary programmes developed from a focus on training into multi-disciplinary assessments of health care functions, organizations and networks. Their standards of assessment have been developed specifically for health care.

While the standards of accreditation are reliable, and the names of accredited hospitals are generally published on individual websites, many hospitals do not participate in voluntary programmes, and criteria and assessment processes vary from program to program. Details of survey results are not publicly available, except for governmental programmes. Measurements of hospitals include internal self-assessment, external survey by multi-disciplinary teams of health professionals, and benchmarking of a limited range of statistical indicators.
A global study\(^1\) identified 36 nation-wide accreditation programmes. A survey of the WHO European Region in 2002 (39) identified 17 such programmes focusing on whole hospitals. Mandatory programmes have recently been adopted in France (40), Italy (41) and Scotland (42).

National programmes within Europe have agreed in principal to voluntary convergence of standards and assessment processes according to the ALPHA Principles of the International Society for Quality in Health Care (43). The ALPHA programme aims to make standards-based assessment systems more reliable, valid and compatible within and between countries (44). Most established programmes have been subjected to internal (45, 46) or external evaluation (47, 48, 49, 50), but few of these evaluations have used comparable methods to permit synthesis. There is ample evidence that hospitals rapidly increase compliance with the published standards and improve organizational processes (51,\(^2\)) in the months prior to external assessment, but there is less evidence that this brings benefits in terms of clinical process and outcome (52,\(^3\)).

The potential for provider profiling from accreditation surveys greatly exceeds what is available from routine statutory returns and minimum data sets, but most accreditation programs do not fully utilize this capacity. Inhibiting factors include the ownership by institutional customers of the raw data, and the costs of developing and maintaining an analytical database without a guaranteed market for its products.

**Statistical indicators**

Statistical indicators can suggest issues for performance management, quality improvement and further scrutiny. They provide relative rather than absolute messages and need to be interpreted with caution inversely proportional to the quality of the underlying data and of the definitions used. Indicators are tools for assessing hospital performance either internally or externally. They should be designed to measure the achievement of predetermined objectives, but in practice they are often selected on the basis of whatever data are routinely available. Standardization is essential for measurements within hospitals, and critical for measurements between hospitals.

Performance measurements from individual hospitals may be submitted as calculated indicators or as raw data to be processed, aggregated, analysed and presented by a central agency. Results are usually disseminated through government publication, website or independent media aimed at consumers, together with guidance on interpretation. Statistical indicators represent an accessible, fairly economical, potentially standard and non-invasive means of performance measurement, but there are many cautions associated with their use: Interpretation of “raw” data on hospital performance, even after adjustment for case-mix and severity, is dependent on many social or economic variables beyond the hospital’s control. Moreover, hospitals might modify internal data collection in order to “meet” external targets, or deny interventions to high-risk individuals in order to improve outcomes. Composite measurements of heterogeneous activity obscure the contribution of their individual elements (53). Many hospitals do not have adequate data to compile standard indicators; the cost of data collection may exceed their value. The time and investment required to develop and validate national indicators are often underestimated. The Sitzia study’s (14) judgement is that “Indicators for the purposes of government inspection and identifying poor providers have had little credibility with providers, and are thought to be unreliable and invalid.”

The European Clearing House on Health Care Outcomes (ECHHO) research project (54) catalogued and linked national initiatives to assess health care outcomes and current projects assessing the results of specific clinical interventions. Another EU research project explored the measurement of appropriateness

---

1 Second European survey of national accreditation programmes. European Accreditation Forum, unpublished.
How can hospital performance be measured and monitored?
WHO Regional Office for Europe’s Health Evidence Network (HEN)
August 2003

of hospital utilization (55). A current OECD project on Health Care Quality Indicators (HCQI) is
developing measures to help decision-makers formulate evidence-based policies to improve the
performance of health systems, rather than of hospitals. In 2003, a WHO Regional Office for Europe
working group (56) began to define performance measures for hospitals’ voluntary self-assessment and
for external benchmarking in six domains: clinical effectiveness, patient centeredness, production
efficiency, safety, staff development and responsive governance. The group has considered background
information on international, national and regional or provincial systems that use standardized data to
evaluate several dimensions of hospital performance for purposes of public reporting, accountability,
accreditation or internal use (57).

According to the Thompson study of clinical indicators (58), “Much of the current evidence on the
effectiveness of performance indicators is based on observational or experiential data, although much of
the policy agenda in this area seems to be based at its worst on dogma.” In any case, it can be said of
statistical indicators usage that:

- In the Czech Republic indicators from routine data showed that, after issuing guidelines to
  standardize management of common conditions, the average length of stay and episode costs were
  reduced without detriment to clinical outcome.
- There are an increasing number of independent reports of the usefulness of some schemes (59).
- A 1995 study in the United Kingdom found that acute myocardial infarction outcome data did not
  show “gross failures of care” (60).
- A 1996 study in the United States showed that outcome data did not identify poor quality hospitals
  (60).
- Research in the USA and Europe has shown wide variations in values expressed by patients, and in
  their use of information designed to empower them (60).
- Published results should highlight broad differences rather than precise rankings (61).

Current debate and trends

Integrating performance measurement systems
Recent national reports from Australia (62), Scotland (63) and the United States (64) have examined how
external mechanisms for performance measurement contribute to internal development and public
accountability. The common conclusions are that:

- Voluntary and statutory agencies should be actively coordinated for consistency and reciprocity.
- Consumers should be prominently involved.
- National programmes should be comparable internationally.
- The standards, processes and results of external assessments should be transparent and freely
  accessible to the public.

Public disclosure of hospital performance data
The publication of hospital activity and results as “league tables” aims to encourage improvement, to
empower patient choice and to demonstrate a commitment to transparency. Evidence suggests that this
increases public interest and management attention to data quality but it does not appear to have much
effect on performance:

- Most publication schemes have been found to have little effect on patient choice behaviour,
  provider behaviour or outcome performance (14).
- The United States Health Care Financing Administration published hospital mortality rates in 1988,
  publication was stopped 1995 because of criticism of the data’s validity and the view that
  publication did not stimulate improvement but caused defensiveness and fear among providers.
• A 1995 survey of Pennsylvania cardiologists found the consumer guide to coronary artery bypass
graft surgery to be “of little or no influence” in choice of surgeon and not much used by consumers
(65).
• One study argues that on statistical grounds, “the current official support for output league tables,
even adjusted, is misplaced” (66).

Conclusions

1. Performance measurement systems should be defined in a published national or regional plan for
quality and performance management that clarifies the values and participation of stakeholders.
2. Governments need to take stock of existing approaches and programmes, to encourage
harmonization of standards, measurements, incentives and public information and to foster
collaboration between the public and private sectors.
3. The underlying values, reference standards and objectives of hospital performance measurement
systems should be made explicit and agreed with stakeholders.
4. The system should not rely on single sources of data but should combine a range of informants.
5. All approaches to performance measurement suffer from behavioural and technical problems, and a
general lack of robust evidence to define their active ingredients.
6. The design of performance measurement systems should aim to manage and improve hospital
performance, rather than to generate unreliable rankings and comparisons.
7. Relevant principles based on international experience include:
   a. Performance failures are more often a result of failures in systems and processes rather
      than of individual competence or knowledge.
   b. Performance assessment requires reliable methods of measurement against validated
      standards.
   c. The reliability of indicators is determined primarily by the accuracy, completeness and
      timeliness of patient-based data collected at institutional level.
   d. Valid comparisons of performance between institutions demand rigorous standardization
      of assessment criteria and methods, especially if they are to be used between countries.

Policy implications

From the viewpoint of policy-making several questions arise, including the following:

• National policy: Is there an explicit, published and comprehensive plan for performance
management and quality improvement? What long-term objectives of that plan is hospital
performance measurement designed to achieve?
• Stakeholder participation: What role would stakeholders (public, professions, insurers, managers)
have in defining, measuring, interpreting and using hospital performance results? Would the same
system be applied to the public and the private hospital sector? How would voluntary performance
measurement be incorporated into a national system?
• Availability, acceptability and credibility: What investment would be necessary and affordable to
provide a complete, accurate and timely common minimum data set for hospitals? What safeguards
and incentives would be available to avoid “gaming” and manipulation of data for political,
financial or commercial reasons?
• Evaluation and publication: In what form would performance data for individual hospitals be
available to stakeholders? What assistance would be provided, especially to the public, on their
interpretation? How would freedom of information be reconciled with confidentiality and data
protection?

The available evidence on hospital performance measurement poses these and other questions, but does
not provide prescriptive answers. Those must be tailored to individual situations, based on collective
How can hospital performance be measured and monitored?
WHO Regional Office for Europe’s Health Evidence Network (HEN)  
August 2003

experience not only in hospitals but in other settings. More details of this experience may be found in WHO Regional Office for Europe’s publication *A background for national quality policies in health systems* (67).
References


How can hospital performance be measured and monitored?
WHO Regional Office for Europe’s Health Evidence Network (HEN)
August 2003


23 National Health Service patients’ survey (http://www.doh.gov.uk).


How can hospital performance be measured and monitored?
WHO Regional Office for Europe’s Health Evidence Network (HEN)
August 2003


How can hospital performance be measured and monitored?
WHO Regional Office for Europe’s Health Evidence Network (HEN)
August 2003

54 European Clearing Houses on Health Care Outcomes (http://www.leeds.ac.uk/nuffield/info/services/ECHHO/home.html).


58 Thomson, R. Clinical indicators: do we know what we are doing? *Quality in Health Care*, 1998, 7:122.


How can hospital performance be measured and monitored?

WHO Regional Office for Europe’s Health Evidence Network (HEN)
August 2003

World Health Organization
Regional Office for Europe
Scherfigsvej 8
DK-2100 Copenhagen Ø
Denmark
Tel.: +45 39 17 17 17. Fax: +45 39 17 18 18. E-mail: postmaster@who.dk
Web site: www.euro.who.int