

SHEA Position Paper

An Approach to the Evaluation of Quality Indicators of the Outcome of Care in Hospitalized Patients, With a Focus on Nosocomial Infection Indicators

The Quality Indicator Study Group

ABSTRACT

The Quality Indicator Study Group was created by the governing boards of three national professional organizations that have interest and experience in epidemiology, nosocomial infection control and prevention, and quality of care improvement. The Study Group has reviewed the existing literature concerning quality indicators (QIs), inter-

viewed experts in the field, and focused on how best to evaluate such indicators, with an emphasis on nosocomial infection indicators as a paradigm for all QIs. In this report, we review pertinent issues and, where possible, provide specific advice on how to evaluate QIs and QI systems (*Infect Control Hosp Epidemiol* 1995;16:308-316).

INTRODUCTION

As the healthcare industry continues to undergo reform-with or without direct governmental intervention-the issues of cost, access, and quality remain in focus. The key question remains: Can access be improved and costs reduced, or at least controlled, while continuing to maintain and improve quality? This important question, while of interest and concern to both providers and purchasers of healthcare, may be most germane to those currently involved in defining and measuring the quality of care.¹⁻⁵

The objectives of this report are to (1) outline epidemiologic principles and propose practical standards for the design and application of healthcare quality indicators; (2) offer thoughtful critique of quality indicators; (3) provide some caveats and cautionary discussion from an epidemiologic perspective

related to current and future quality indicators and their interpretation and use by end users; and (4) propose a responsible approach to prevent wasted effort and expenditure of healthcare dollars due to hastily employed, ill-crafted approaches to quality indicator development. Such approaches, good intent notwithstanding, carry the risk of generating misleading feedback information. The potential for subsequently flawed decision making can only compound an already complex situation.

This review aims to provide a comprehensive discussion of the above issues for several potential groups of consumers. The primary audience of this work includes those directly involved in healthcare quality management, especially those practicing infection control and hospital epidemiology. This article is intended to enlighten them concerning the issues discussed herein, as well as to provide a helpful

*From the Quality Indicator Study Group, the Society for Healthcare Epidemiology of America: *William E. Scheckler, MD, Chair (Department of Family Medicine, University of Wisconsin Medical School, Madison, Wisconsin); *Robert Gaynes, MD (Hospital Infections Program, Centers for Disease Control and Prevention, Atlanta, Georgia); *Peter Gross, MD (Department of Internal Medicine, Hackensack Medical Center, Hackensack, New Jersey); *Walter Hierholzer, Jr, MD (Department of Epidemiology, Yale-New Haven Hospital, New Haven, Connecticut); *Robert A. Weinstein, MD, (Division of Infectious Diseases, Cook County Hospital, Chicago, Illinois). From the Association for Professionals in Infection Control and Epidemiology: *Ona Baker, RN, MSHA, CIC (Northwest Texas Healthcare System, Amarillo, TX); Jacalyn Bryan, RN, BSN, CIC (Georgetown University School of Nursing, Washington, DC) [until 2/1/94]; *Terrie Lee, RN, MS, MPH, CIC (Charleston Area Medical Center, Charleston, West Virginia); *Emily Rhinehart, RN, BSN, CIC (AIG Consultants Inc., Boston, Massachusetts) [after 2/1/94]. From the Surgical Infection Society: *James T. Lee, MD, PhD (Veterans Administration Medical Center, Minneapolis, Minnesota). *Indicates authors of report.*

Address reprint requests to the Society for Healthcare Epidemiology of America, 875 Kings Highway, Woodbury, NJ 08096-3172. 95-SR020. The Quality Indicator Study Group. An approach to the evaluation of quality indicators of the outcome of care in hospitalized patients, with a focus on nosocomial infection indicators. Infect Control Hosp Epidemiol 1995;16:308-316.

TABLE 1
SHEA STATE LIAISON REPORTS OF STATEWIDE QUALITY INDICATOR SYSTEMS IN USE OR BEING PROMOTED

State	Indicator Source and Promoter	Optional or Required	Includes Infection	Nosocomial Indicators
California	University of California Davis Medical Center	Optional		Yes
Connecticut	Community Hospital Association-Maryland Hospital Association Quality Indicator Project (MHAQIP)	Optional		Yes
Georgia	Georgia Hospital Association	Optional		Yes
Indiana	Indiana Hospital Association	Optional		Yes
Maryland	MHAQIP	Optional		Yes
Massachusetts	43 Hospitals (MHAQIP)	Optional		Yes
New Hampshire	New Hampshire Hospital Association	Optional		Yes
New York	PRO Medical	Optional		Yes
North Dakota	North Dakota Hospital Association (MHAQIP)	Optional		Yes
Ohio	State Government	In development-anticipate required		Yes
Rhode Island	RI Hospital Association (MHAQIP)	Information not available		Yes
West Virginia	West Virginia State Hospital Association (MHAQIP)	Optional		Yes
Wisconsin	Wisconsin Hospital Association (MHAQIP)	Optional		Yes

resource when faced with decisions related to selection of quality indicators or to participation with other organizations in the collection, aggregation, analysis, and use of data.

A second audience for this work includes decision makers and senior managers in acute-care institutions who have been, and will continue to be, solicited for the participation of their organizations in voluntary quality indicator projects. The discussion provided also should facilitate more knowledgeable participation in mandatory quality indicator programs, such as that proposed by the Joint Commission on Accreditation of Healthcare Organizations.

The providers of quality data must be fully knowledgeable of the epidemiologic principles that apply to these activities, as well as the many perils and caveats that may result if these principles are ignored or not fully employed in the design, collection, aggregation, and analysis of data related to quality indicators.

Finally, another audience for this work may be the purchasers and receivers of healthcare. To date, they too have remained unsatisfied in the quest to define quality healthcare. This information may provide them with new insights into future methods to measure quality of care.

CURRENT STATUS OF UTILIZATION OF QUALITY INDICATORS

In October 1994, a survey of statewide indicator systems was sent to the Society for Healthcare Epidemiology of America (SHEA) state liaisons in the 48

states where at least one liaison was present. Responses were received from representatives in 28 states. Liaisons from the following 15 states indicated that no formal statewide quality indicator (QI) system was planned or in use: Alabama, Arizona, Arkansas, Colorado, Florida, Kentucky, Maine, Mississippi, Nebraska, North Carolina, Oregon, Tennessee, Texas, Vermont, and Washington. Liaisons from another 13 states indicated that some type of statewide quality indicator system was being promoted (Table 1).

The speed of the development of QIs and their implementation has accelerated. Two organizations have been involved extensively in development or provision of QIs for use in acute-care hospitals: the Joint Commission on Accreditation of Healthcare Organizations (JCAHO) and the Maryland Hospital Association Quality Indicator Project (MHAQIP). The JCAHO began the development of its Indicator Measurement System as part of the Agenda for Change in 1989. Indicators have been developed, field tested, published, and approved for voluntary use in several disciplines, including anesthesia, obstetrics, cardiovascular surgery, and trauma. Indicators are in development for medication use, infection control, and depression. In order to provide data for comparison and accreditation of hospitals, the JCAHO may require implementation of selected indicators by all accredited hospitals as soon as 1997.

The MHAQIP had approximately 950 hospitals participating in their QI project as of late 1994. Participation is coordinated through state hospital associations and other hospital networks. Participat-

ing hospitals collect and provide their data through their sponsoring organization. Aggregate data are produced and provided to participants for their internal use.⁶ The MHAQIP includes several nosocomial infection indicators, which MHAQIP currently is reviewing.

The National Nosocomial Infection Surveillance (NNIS) System of the Centers for Disease Control and Prevention (CDC) has been developing and refining measures of the incidence of nosocomial infections since 1969. Insofar as nosocomial infections are seen as a QI measure, this system has the most extensive experience with QIs. NNIS currently monitors site-specific nosocomial infection rates in patients at risk in intensive care units (eg, ventilator-associated pneumonias) and surgical site infection rates, stratified by a risk index that is tailored to the population of patients undergoing a specific surgical procedure. The current NNIS methodology is a useful benchmark against which nosocomial infection quality indicators from other sponsors can be compared.⁷

This report expands and refines our interim report⁸ and represents a multiauthor collaboration by doctors and nurses who currently perform epidemiologic research, provide direct patient care, and/or manage clinical epidemiology and quality improvement programs. We have tried to keep Tukey's dictum in mind: *approximate answers to right questions are worth more than precise answers to wrong questions*⁸. Our consensus views and corresponding practical advice have evolved with the final goal of crisply distinguishing areas of clear understanding from areas of uncertainty. This report has been approved by the governing boards of the Society for Healthcare Epidemiology of America, the Association for Professionals in Infection Control and Epidemiology, and the Surgical Infection Society, and is being published simultaneously in *Infection Control and Hospital Epidemiology* and the *American Journal of Infection Control*.

IMPORTANT ISSUES IN EVALUATING QUALITY INDICATORS

Use of Outcome or Process Indicators

Most organizations attempting to examine quality in the delivery of health care include an analysis of a variety of events that occur in the hospital, such as nosocomial infections. As institutions around the country aggregate surveillance data on nosocomial infection rates, the use of these data for interhospital comparisons has become the subject of debate among hospital epidemiologists and infection control professionals.⁷ A hospital must consider carefully its participation in one of the various multihospital surveillance systems. The decision to participate should be dis-

cussed at all levels within the hospital—from the hospitals senior management to the persons involved in collecting the data. There must be a link between the collection of such data and a continuous improvement strategy so that the caregivers can improve the quality of health care. In addition, the evaluation of any surveillance system for quality improvement must be based on detailed review of the factors discussed below.

Some of the events that individuals or groups have used as indicators of quality and that reflect outcomes are actual patient complications of hospitalization, eg, nosocomial infection rates. Others are events that are not related directly to an outcome per se, eg, urinary catheterization rates. The latter, called process indicators, do not measure a patient outcome, but nonetheless can be useful to evaluate quality if—and only if—they can be linked to an outcome measure, eg, urinary tract infections. Process indicators also are worthwhile if the outcome to be assessed is rare; if cost containment is of concern; if used as a measure of process of an educational effort in hospitals; or if used as a measure of behavior (eg, compliance with isolation policy), even when scientific proof of efficacy may be lacking.

Selection of Quality Indicators

Is an indicator an accurate measure of quality? This question is difficult, if not impossible, to answer until we have a universally accepted definition of quality health care. However, there is active evaluation of various indicators. For example, there are many sponsors of indicator systems that collect nosocomial infection rate indicators (as reported before⁸ and noted above). Each of the rates is being evaluated as a measure of quality by most of these sponsors. Perhaps the best approach would be to ask the question: As you improve the indicator (eg, decrease the occurrence of adverse events such as nosocomial infections), do you improve patient outcome? Note that indicator rates may change because of changes unrelated to changes in quality, such as a change in the case mix or in the system of case finding.

Definition of an Indicator—The Numerator

The numerator is the event being tracked (for example, a patient fall or wound infection). The selection of specific nosocomial infections or events as indicators usually should focus on those (a) that have clear cut definitions, are supported by prior studies, and can be applied readily in most hospitals; (b) that are readily identified; (c) that are clinically important in a given patient population; and (d) that are amenable to risk stratification.⁷ No indicator will be perfectly appropriate for every institution or for every patient

TABLE 2
COMPARISON OF FACTORS IMPORTANT IN SELECTING INDICATOR EVENTS: EXAMPLES OF NOSOCOMIAL INFECTION INDICATORS IN ADULTS

Factors	Pneumonia*	Urinary Tract	Surgical Site	Bacteremia
Clear case definition	2+	4+	3+	4+
Ease of specimen collection	2+	4+	3+	4+
Ease of surveillance				
Laboratory	2+	4+	2+	4+
Clinical	2+	2+	3+	2+
Relative frequency of event	3+	4+	3+	2+
Importance of event				
Morbidity	3+	2+	3+	4+
Mortality	3+	1+	2+	4+
Potential of interventions to reduce rates	2+	3+	3+	3+
Ease of stratification				
Exposure risk factors	3+	4+	4+	4+
Severity of illness risk factors	3+	2+	3+	2+
Denominator by device days				
Availability	Y	Y	N/A	Y
Ease of collection	V	V	N/A	V

Abbreviations: Y, yes; N/A, not applicable; V, varies with setting.

*Factors are ranked from 1+ (least favorable) to 4+ (most favorable).

population. It is important to avoid overly broad indicators, such as "all nosocomial infections." Moreover, some important indicators may be difficult to measure precisely. For example, despite the importance of nosocomial pneumonia as a cause of hospital morbidity, the clinical case definition of nosocomial pneumonia is neither sensitive nor specific; conversely, nosocomial urinary tract infections are easily defined and detected¹⁰ but infrequently are important factors in patient morbidity or mortality. Table 2 compares several indicators with respect to characteristics that should be considered when choosing specific nosocomial infections as indicator events.

Definition of an Indicator--The Denominator

The denominator is the patient population at risk of experiencing the event that appears in the numerator. The selection of denominators for the indicator events (and definition of a time period) allows the development of indicator rates (numerator/denominator = rate). Two important decisions in developing or selecting a denominator are how specific to be (eg, rate per 100 patient discharges is less specific than rate per 1,000 patient days, and rate per 1,000 device days may be the most specific) and how to stratify patients by their risk factors. Table 3 lists examples of denominators and of approaches to stratification by risk.

Because rates may be tracked over time and compared among wards, services, and institutions, it is important to use specific denominators and to stratify (separate into homogeneous subgroups) the data. Usually, the more specific the stratification, the more meaningful the comparison. For example, surgical site infection rates can be divided into groupings, or strata, that account for differences in infection risk. It has been found that stratifying on a composite risk score that considers general physical status (for example, the American Society of Anesthesiologists preoperative classification score), wound contamination classification, and the length of the procedure produces more meaningful comparisons than does lumping all wound infection numerators together or stratifying on wound class alone.¹³

Reliability of Data Collection

To be useful for interpretation, QIs must generate reliable and reproducible results. Reliability may be defined as the degree to which the QI correctly identifies the event targeted for monitoring. Reproducibility evaluates whether the findings based on the use of given QIs can be repeated consistently when applied to new populations, to different institutions, or by different individuals. Likewise, methods should demonstrate a high degree of interrater reliability: two raters (surveillance staff) reviewing the same set of medical information should consistently classify events (such as infections) in the same way.

Definitions and data collection procedures should be in writing, with sufficient detail to foster clarification of methods, whether for one or numerous data collectors.^{10, 14, 15} Operationally, the emphasis should be placed on training in the use of valid, clearly written descriptions of data elements and the strict adherence to those criteria.

In addition to limitations in training and quality control methods, another impediment to reliability may be the medical model itself. It is not uncommon for data collection staff to defer to the opinions of a clinician about the presence or absence of an infection or other event, rather than simply to determine whether case definitions are met. This inclination to make decisions on a case-by-case basis is consistent with the medical model of individualized care and the peer review process, but not with the epidemiologic model of population-based analyses. Clear distinctions between case definitions for surveillance purposes and case definitions for clinical diagnoses and treatment are crucial.

The limitations of small sample sizes also must be considered carefully. While the data collected may be consistently accurate, analyses and subsequent conclusions may be constrained by the small numbers studied. Organizations with small populations are intrinsically limited, and few guidelines exist regarding how to address the evaluation of QIs in such settings. It has been proposed that studies of validated process measures having documented association with specific outcomes might be most appropriate when dealing with infrequent outcomes. Even when dealing with large numbers of cases, risk stratification methods necessary for valid comparisons often result in small numbers in some risk strata, thus limiting intrahospital or interhospital comparisons.

Completeness of Data Collection

Just as the reliability and accuracy of the collected data are key factors in the utility of QIs, the completeness of data collection is a similar concern. There should be explicit written descriptions of data collection methods that minimize ad hoc decisions. Surveillance intensity must be consistent if data are to be compared over time or among institutions. Data collection can be compromised by suboptimal staffing or uncovered absences of the data collection staff, treating data collection as a low priority activity, inadequate computer resources, and incomplete chart information or availability.

Additionally, methods should be in place to verify the completeness of the ongoing monitoring. These would be linked to reliability determination procedures and should be formalized and performed routinely. If healthcare systems, consortia, or other

TABLE 3
EXAMPLES OF DENOMINATORS FOR INDICATOR EVENTS

Hierarchical List of Denominators (most general to most focused)
Patient admissions or discharges
Patient hospital days
Patient days by specific units
Patients exposed to a specific device
Patient days of exposure to a specific device
Patient days of exposure to a specific device in a specific risk group
Examples of Stratification by Risk
Device Exposure
Demographic factors
Underlying disease
Other infection risk factors (eg, medication exposure)
Physiologic factors (eg, APACHE score)
Other severity of illness measures (eg, wound classification)
Composite (multivariate) indices considering multiple risk factors

bodies aggregate data from multiple institutions, these data quality monitors are mandatory, and commitment to them should be part of the initial enrollment prerequisites. Without validation of completeness and accuracy of data, comparisons by aggregating systems probably are meaningless and, potentially, very misleading.

Measures should be incorporated to verify the accuracy of data entry into automated databases. This could be done by overreading of a sample of records input by other staff, by appropriate error checks in the software, and by methods to determine data integrity routinely. Case vignettes have been used to evaluate the accuracy of use of standard definitions of nosocomial infections.¹⁶

Finally, experts in the field should participate in the development or validation of data collection methods used by all aggregating systems. Descriptions of training methods should be published for review and should be standardized wherever possible.^{16, 17}

Training

Implementation of a quality indicator system should include not only formal training of those who will be collecting, analyzing, and communicating data, but also appropriate training of the senior management, trustees, medical staff, and managers. This can enable an institution to maintain an environment that will support the implementation of a monitoring system. When senior managers are not aware of the

implications of the quality indicator system, including the resources required for its implementation and the appropriate use of data for comparative analyses, barriers to successful implementation and interpretation are likely to be encountered.

Ideally, training of data collection personnel should include the opportunity to practice techniques such as the application of case definitions, as well as to provide feedback on performance.^{6,7} This can be a critical aspect of ensuring the reliability and reproducibility of future data. Methods of appropriate training range from formal courses with practice workshops to self-instructional methods, such as computer-assisted instruction,¹⁷ to mentoring. Key to the success of training is the involvement of experts from the related field (eg, for infection control indicators, experts in infection control and infectious diseases), as well as epidemiologists and other process experts. Effectiveness of models for training trainers to deliver educational messages to broad groups will be determined largely by the level of expertise of the trainers.

Feasibility and Ease of Data Collection

While the ideal data set for intrahospital or interhospital comparisons of clinical outcomes could include an exhaustive list of potential risk factors and their important parameters, the size of the list must be manageable for a hospital's resources. For broad application, quality indicator data elements must be amenable to accurate and consistent collection. Benefits gained or lost by adding or deleting specific elements must be weighed carefully. The evaluation of benefit should reflect the likelihood that the additional information will assist in clinical decision making and, ultimately, in improvement of outcomes. Limited resources ordinarily would be directed toward issues (1) for which documentation exists that correlates data collection and reporting with improved outcomes,¹⁸ (2) having the greatest theoretic potential for improvement, or (3) having the most grave consequences. Human resources (staffing needs) also should be considered explicitly. The development of contemporary guidelines for staff allocation currently is a critical research item among infection control professionals.

Comparability of Populations: Severity and Case Mix Realities

When assessing a healthcare process or outcome indicator, adjustment for case mix and severity is important.⁸ Comparison of the performance of healthcare practitioners or institutions may be misleading without adjustment for these (and perhaps others, such as practice patterns). However, the comparability among institutions of indicator results that have been severity adjusted cannot be assumed. Severity of

illness indicators must be used with proper statistical methodology or very misleading results can occur, as a group from Pennsylvania recently demonstrated.⁴

The level at which to begin the comparison also is critical. If the level is too coarse, the heterogeneity in the case mix may be too great for meaningful comparison. If the level is too refined, the number of cases available for comparison is likely to be small, compromising the statistical power of the comparison. Within the DRG system, a gradient of homogeneity exists. The progression from the least to the most homogeneous is as follows: MDC (Major Diagnostic Category) → DRG (Diagnosis-Related Group) → ICD-SCM (International Classification of Diseases, 9th Revision, Clinical Modification) → severity system score.²⁰ The next revision of the DRGs likely will include adjustment for severity of illness. The Health Care Financing Administration is considering the Yale Refined DRGs and the New York All Patient DRG systems.

Comparison at the DRG level usually is chosen by most institutions in an attempt to yield meaningful numbers, but the reliability of comparison at this level may not be great if the case mix and severity vary significantly among the institutions being compared. As an example, consider mortality from pneumonia at two hospitals. Analysis could begin with all patients in DRG 89 (simple pneumonia and pleurisy: age >17 with complications or comorbidity) and DRG 90 (simple pneumonia and pleurisy: age >17 without complications or comorbidity). But DRGs 89 and 90 include at least 15 different diagnostic codes, such as ICD-SCM 481 (pneumococcal pneumonia), 482.9 (unspecified bacterial pneumonia), and 487.0 (influenza with pneumonia). One community hospital may have a preponderance of pneumococcal pneumonia patients (with low severity of illness scores) in the DRG, while the other hospital may have a large organ transplant service and many patients with unspecified bacterial pneumonia (and high severity of illness scores).

Available systems for assessing severity have different forms and uses.¹⁹ Some are intended for almost all diseases, such as APACHE II or III or Medis-Groups.^{20,21} Others are disease specific, such as the Ranson Criteria for mortality from pancreatitis or the Dukes staging system for 5-year survival following colorectal cancer.^{22,23} Still others are procedure specific, such as the NNIS risk index for surgical site infections.²⁴ Validation of the use of generic severity indices for nosocomial infections continues to evolve.

Internal Tracking Versus External Comparisons: Benchmarks or Standards of Care

It has been shown that an indicator that is appropriate for comparing healthcare practitioners within an institution nonetheless may be inappropri-

ate for comparing results at two or more institutions.⁷ Therefore, an indicator always should be validated for any new use, particularly for comparisons. Indeed, concern for the validity of such comparisons prompted formation of the QI Study Group.

Before using a reported indicator rate as a benchmark, it is critical to know whether the setting and data collection techniques are comparable. For example, was surveillance at the benchmarking institution comparable to the other institutions? Are the definitions of disease and risk comparable? Furthermore, true benchmarking requires knowing who is "best in class" for a given outcome or process, with efforts made to top that result. It means knowing who the "best" is so that this entity can be used as a resource for learning how to improve care. Most comparative databases do not identify the reporting institutions, so that true benchmarking is not possible. Usually, results at an organization simply are compared to the range of results reported by similar institutions. Large severity-adjusted databases have been gathered by commercial healthcare organizations.^{20,25-27} Their validation by independent investigators is just beginning.

Appropriate statistical testing should be part of both internal and external comparison exercises. This is critical especially when comparative data are used in the process of credentialing and privileging individual practitioners or in institutional accreditation. Further, accurate analysis of data can ensure that resources are directed toward clinically important issues.

Confidentiality Versus Marketing

The concept of measuring quality has been mentioned frequently in this era of healthcare reform.^{28,29} Parties interested in these measurements (often posed as "quality report cards") include managed care plans, accrediting or licensing agencies, insurance companies, employers, and individual consumers. These entities may require reporting of quality data, including infection rates, thus forcing participation in an aggregate database. It is important to identify how the aggregating body intends to use the data and to create opportunities for educating appropriate parties about interpretation of data, limitations of comparisons, and other supporting issues discussed here. If misuse of data is likely to occur, this activity becomes even more important. When one is asked to supply inappropriate data (eg, crude infection rates that are not adjusted for risk), one might consider reporting only accurate, adjusted rates and offering to assist in the educational process as suggested above. It is essential for hospitals and healthcare providers to insist on wording in contracts that specifies the proper epidemiologic attrib-

utes before outcome data would be provided by the hospital.⁷

Relation to Continuous Quality Improvement CQI

Accurate indicator data, including infection rates, may be useful and supportive of CQI efforts within an organization. In the CQI paradigm, clinical decisions are based on data that are gathered and analyzed over time, with the use of appropriate statistical tools and with feedback given to clinicians. CQI efforts can address both endemic levels of disease (as common cause variation), as well as epidemic levels of disease (as special cause variation). Good epidemiology is as important for use of data internally as it is for external comparisons.

Indicator systems can measure outcomes, such as infections; in addition, related processes also are of interest in the CQI world. Multidisciplinary teams can address the multivariate nature of the risk of infections in an effort to improve both processes and their associated outcomes.³⁰⁻³⁵

Timeliness of Data Collection and Reports

Data are most useful when the time between data collection and availability for review is short. In terms of intrahospital use, rapid analysis of data is key to the ability to identify opportunities for improvement. Generation of timely internal reports is important to the function of supplying data perceived to be useful by clinicians. Also important is the ability of the individual organization to generate their own analyses from data collected for the indicator system; this may require that centrally distributed software have ad hoc reporting capabilities or, at the least, generate data files of a known and standard format. The rapid reporting of data to an aggregating agency can assist in ensuring completeness of the overall database.

If interhospital comparisons are to be used, reports from the aggregating body should be timely for analyses to be useful in the improvement of performance. Delays in reporting diminish the positive nature of feedback.

Sharing Methods and Data in the Medical Literature

In the United States, many organizations have supported the development of guideline and indicator outcome activities. Differing methodologies and nomenclatures have been employed. Some have been specifically for or related to infection control. To avoid duplication and increase effectiveness and the timely implementation of the use of indicators and guidelines, it is important to share methodologies, analyses, and results across a wide population of users.

The Institute of Medicine of the National Academy of Science published a monograph in 1990 reviewing the development and use of clinical practice guidelines.³⁶ The report listed desirable attributes of guidelines and the medical review criteria used in their development. These qualities may be considered as equally important features of indicators. They include validity, reproducibility, clinical applicability, flexibility, clarity, sensitivity, specificity, feasibility, and documentation of the methods used in development and of the evidence used for assumptions.

Other authors and groups have commented on topics that could serve as the basis for development of indicators, the usual goals of indicators, and recommendations for guideline or indicator development.^{1,5} As in all practice analyses, it is important to choose a right question. The Agency for Healthcare Policy and Research has suggested guideline selection based on (1) adequacy of the supporting science, (2) the population size affected, (3) the potential for prevention or improvement, (4) the cost of the condition, and (5) the associated needs of other government programs. While the goals of a guideline usually will be to improve quality, other reasonable ends might be to control cost, to provide guidance in ethical or legal decisions, or to improve effectiveness of a process.

Guidelines and indicators may be developed by many methods. Investigators at the Rand Corporation have suggested (1) an in-depth literature review with associated critiques of the quality of the literature as part of the process, (2) presentation of the findings of the review to a panel of experts with development of a consensus by some variation of delphi techniques, and (3) the association of the recommendations with a rating system indicating to the user the strength of the research and opinion support for components of the recommendation(s).

Once completed, indicators must be pilot- and field-tested. Pilot testing may be through retrospective cohort studies, but subsequent field testing (the so-called alpha- and beta-testing) must be accomplished under practice conditions. Results of these analyses should be shared with the expert development group, and corrections should be made to improve efficacy and efficiency. Many indicators and guidelines do not perform well outside the artificial development environment and fail at this point. Indicators surviving these tests are ready for release.

Successfully developed indicators must be disseminated widely to the appropriate users. Presentation of indicators in peer-reviewed scientific journals respected by the targeted users is one method of assuring another level of quality review and broad dissemination to the practicing field.

CONCLUSION

This is the final report of the joint Quality Indicator Study Group, which was established in January 1993 by the participating organizations to perform the following tasks:

1. Review quality indicators that are in use or in development and that are or will be used for interhospital comparisons of the outcome of care among acute-care hospitals in the United States;
2. Describe scientific and epidemiologic criteria that may be used to evaluate these quality indicators, with a focus on outcome indicators for nosocomial infection;
3. Apply these criteria to existing and planned quality indicators to prepare a state-of-the-art position paper for publication in the journals of the study group participants.

This paper has focused on the second item above, because the group felt this would be the most useful part of the exercise. Existing indicator systems are in such a state of flux or development that producing a table comparing current systems with our criteria would be premature. (The NNIS system, however, is well developed, as noted in the bibliography.)

Quality indicator "report cards" for evaluation of outpatient care and managed care systems also are being developed. A good example is the "Health Plan Employer Data Information Set" (HEDIS).³⁷ The study group did not review these indicator sets, but it is likely that all of the issues and caveats raised above for inpatient QIs also are applicable to outpatient QIs.

Throughout our deliberations, consultations with groups, discussions with members of our organizations and our boards, review of the literature, and our production of this report, it has been our goal to educate ourselves so that we might be more helpful to our membership. To the extent that our interim report and this article are helpful in improving the science and rationality of the use of quality indicators, we have succeeded. As should be clear from this report, however, much remains to be done.

REFERENCES

1. Public accountability of hospitals regarding quality: future directions; report of an invitational workshop sponsored by AHCPR; December 3, 1993; Philadelphia, PA. The joint Commission *24mal on Quality Improvement* 1994;20:7.
2. Braun BI, Carr MP, Cheng AM. Hospital attributes associated with successful participation in a comparative performance measurement data base. *Clinical Performance and Quality Healthcare* 1994;2:242-250.
3. Berwick D, Godfrey A, Roessner J. *Curing Health Care: New Strategies for Quality Improvement*. San Francisco, CA: Jossey-Bass; 1990.
4. Localio AR, Hamory BH, Sharp TJ, et al. Comparing hospital mortality in adult patients with pneumonia: a case study of statistical methods in a managed care program. *Ann Intern Med* 1995;122:125-132.

5. *Health Care Reform-Report Cards' Are Useful but Significant Issues Need To Be Addressed*. Report to the Committee on Labor and Human Resources, US Senate, Washington, DC: US General Accounting Office; September 1994. GAO publication No. HEH-94-219.
6. Vibbert S, Strickland D, Migdail K, et al. Major projects—Maryland quality indicator project. 1995 ed. In: *Medical Outcomes and Guidelines Sourcebook*. New York, NY: Faulkner and Gray; 1994:29-31.
7. Gaynes R. Nosocomial infection rates for inter-hospital comparison: limitations and possible solutions. *Infect Control Hosp Epidemiol* 1991;12:609-621.
8. Scheckler WE. Interim report of the quality indicator study group. *Infect Control Hosp Epidemiol* 1994;15(4):265-268.
9. Chatfield C. *Problem Solving A Statisticians's Guide*. London, England: Chapman-Hall; 1988:8.
10. Gamer JS, Jarvis WR, Emori TG, Horan TC, Hughes JM. CDC definitions for nosocomial infections. *Am J Infect Control* 1988;16:128-140.
11. New Classification of Physical Status. *Anesthesiology* 1963;24:111.
12. Owens WD, Felts JA, Spitznagel EL Jr. ASA physical status classification: a study of consistency of ratings. *Anesthesiology* 1978;49:239-243.
13. Culver DH, Horan TC, Gaynes Rp, et al. Surgical wound infection rates by wound class, operation, and risk index in U.S. hospitals, 1986-1990. *Am J Med* 1991; 91(suppl 3B):152S-157S.
14. Emori TG, Culver CH, Horan TC, et al. National Nosocomial Infections Surveillance (NNIS) system: description of surveillance methodology. *Am J Infect Control* 1990;19:19-35.
15. Emori TG, Gaynes RF. An overview of nosocomial infections including the role of microbiology laboratory *Clin Microbiol Rev* 1993;6:428-442.
16. Larson E, Horan T, Cooper B, Kotilainen HR, Landry S, Terry B. Study of the definition of nosocomial infections (SDNI). *Am J Infect Control* 1991;19:259-267.
17. Lancaster D, Willis MA. Computer-assisted instruction (CAI): a time-saving, individualized teaching methodology. *Am J Infect Control* 1994;22:179-81.
18. Haley RW, Culver DH, White JW, et al. The efficacy of infection surveillance and control programs in preventing nosocomial infections in U.S. hospitals. *Am J Epidemiol* 1985;121:182-205.
19. Gross PA. Use of severity of illness indices. In: Mayhall G, ed. *Hospital Epidemiology and Infection Control*. Baltimore, MD: Williams&Wilkins. In press.
20. Knaus WA, Douglas PW, Draper EA, et al. Clinical investigations in critical care: The APACHE III prognostic system. Risk prediction of hospital mortality for critically ill hospitalized adults. *Chest* 1991;100:1619-1636.
21. Iezzoni LI, Hotchkin EK, Ash AS, Schwartz M, Mackieman Y. MedisGroups databases: the impact of data collection guidelines on predicting in-hospital mortality. *Med Care*. 1993;31:277-283.
22. Ranson JHC, Rifkind KM, Roses DE, Fink SD, Eng K, Spencer FC. Prognostic signs and the role of operative management in acute pancreatitis. *Surg Gynecol Obstet* 1974;139:69-81.
23. Mayer RJ, Rubenstein E, Federman DD, eds. *Scientific American Medicine: Gastrointestinal Cancer*. New York, NY: Scientific American Medicine Inc; 1986;VIII:1-18.
24. Horan TC, Gaynes RF, Mortone WR, Jarvis WG, Emori TG. CDC definition of nosocomial surgical site infections, 1992: a modification of CDC definitions of surgical wound infections. *Infect Control Hosp Epidemiol* 1992;13:606-608.
25. *The Guide to Hospital Performance*. Baltimore, MD: HCIA Inc.; 1993.
26. Fine MJ, Singer DE, Hanusa BH, Lave JR, Kapoor WN. Validation of a pneumonia prognostic index using the MedisGroups Comparative Hospital Database. *Am J Med* 1993;94:153-159.
27. Acuity Index Method database. Iameter, San Mateo, CA.
28. Berwick DM. The clinical process and the quality process. *Qual Manage Health Care* 1992;1:1-8.
29. O'Leary DS. The measurement mandate: report card day is coming. *J Qual Improvement* 1993;19:487-491.
30. Kritchevsky SB, Simmons BP. Continuous quality improvement. Concepts and applications for physician care. *JAMA* 1991;266:1817-1823.
31. Kritchevsky SB, Simmons BP. Variation in medical practice: implications for the hospital epidemiologist. *Infect Control Hosp Epidemiol* 1994;15:116-119.
32. Laffel G, Blumenthal D. The case for using industrial quality management science in health care organizations. *JAMA* 1989;262:2869-2873.
33. Merry MD. Total quality management for physicians: translating the new paradigm. *QRB Qual Rev Bull* 1990;16:101-105.
34. Scheckler WE. Continuous quality improvement in a hospital system: implications for hospital epidemiology. *Infect Control Hosp Epidemiol* 1992;13:288-292.
35. Sellick JA Jr. The use of statistical process control charts in hospital epidemiology. *Infect Control Hosp Epidemiol* 1993; 14:649-656.
36. Institute of Medicine. *Clinical Practice Guidelines: Directions for a New Program*. Committee to Advise the Public Health Service on Clinical Practice Guidelines. Field MJ, Lohr KN, eds. Washington, DC: National Academy Press; September 1990.
37. Kenkel PJ. Health plans face pressure to find 'report card criteria that will make the grade. *Modern Healthcare* January 10, 1994;24(2):41-42.

Joint Commission Announces Action Plan to Revise Accreditation Process

by Gina Pugliese, RN, MS
Medical News Editor

In response to concerns raised by the American Hospital Association, the Board of Commissioners of the Joint Commission on Accreditation of Healthcare Organizations approved an action plan. In doing so, the Joint Commission affirmed its commitment to high quality patient care and emphasized the Joint Commission's intent to improve and strengthen its accreditation service to support that objective.

The plan identifies specific

improvement initiatives and an ambitious timetable for completion. Some of the improvements include responsiveness to telephone calls and correspondence; a price freeze on fees charged for regular full surveys; a substantial reduction in the number of focus surveys; mechanisms to identify variations in surveyor performance; revision of the survey process to include daily meetings to replace the exit conference; simplification of standards to focus on issues related to quality of patient care; and a review of educational programs and publications

for quality and internal consistency. Finally, the Joint Commission promised to reconfigure the Indicator Measurement System to incorporate measures and measurement systems developed by other entities, with the long-term objective of offering a varied menu of relevant measures to all types of organizations as part of the accreditation process.

FROM: Statement by the Joint Commission of Accreditation of Healthcare Organizations; January 23, 1995.