

# I. Introduction / Overview

## A. Movements Toward Organized Delivery Systems

Continuing cost pressures from purchasers, excess capacity in inpatient and specialty areas, and a desire to create higher levels of clinical integration have created an environment favorable to the creation of organized healthcare delivery systems.<sup>1-3</sup> These systems may be organized and managed by a single health plan or may serve as a provider for several health plans.<sup>4</sup> They come under many different labels and structures—staff model HMOs, multispecialty group practices, Physician Hospital Organizations (PHOs), healthcare networks, Physician Sponsored Organizations (PSOs), regional hospital systems, or simply “health systems.” In spite of the variation in structure and names, certain key concepts are common to most of these organizations. They include:

- service to a defined geographic area;
- service to a defined patient or member population;
- ownership or active management of a wide spectrum of clinical services, including inpatient, outpatient specialty, physician office, rehabilitation, and home health services;
- internal data management capabilities that allow analysis of and influence over clinical care patterns;
- some ability and willingness to accept financial risk (e.g., capitation payment or other forms of bundled payment);
- ability and willingness to report on various measures of quality of care.

## B. Concepts of Accountability and “Value Purchasing”

Public and private purchasers who deal with these organizations have typically requested information about quality in requests for proposals or as part of an employee open enrollment process. This information is requested to provide some assurance to the purchaser that: (a) the provider organization is formally accountable for some defined aspects of quality of care; and (b) purchasers can make informed decisions among competing organizations on the basis of cost, quality, and a combination of the two concepts that might be termed “value.”<sup>5,6</sup>

The concepts of accountability and value shift over time as the relationships among purchasers, health plans, and providers change. Until fairly recently, purchasers were generally satisfied to contract with plans and providers who had received accreditation by the appropriate agencies and presume that quality was at least acceptable and probably not significantly different among competing alternatives. However, the small-area variation studies of John Wennberg and colleagues,<sup>7-9</sup> and the RAND studies of appropriateness of surgical procedures,<sup>10,11</sup> created an atmosphere of “doubt” in the mid to late 1980s about the inherent quality provided by licensed, board-certified physicians and accredited hospitals. The Wennberg studies indicated wide variability in practice patterns from city to city or county to county, with no obvious explanation in population health characteristics and no obvious association with measurable outcomes. The RAND studies indicated that as many as one-third of surgical procedures of various kinds were done without appropriate indications or obvious benefit to the patient.

With these observations in hand, Paul Ellwood issued a public call for a national program of “Outcomes Management” in 1988,<sup>12</sup> and Congress established the Agency for Health Care Policy and Research in 1989.<sup>13</sup> Both efforts represented an attempt to find out “what works” in healthcare in order to develop guidelines for more effective care. Large private employers, most notably Xerox, GTE, and Honeywell, began to define sets of quality measures or quality indicators to be used in judging the care provided by competing health plans.<sup>14</sup> At the same time, the Health Care Financing Administration (HCFA) had begun issuing a national report of hospital mortality rates in order to identify for the public hospitals with either better or poorer mortality compared to the expected rate given the hospital’s case mix. We moved fairly quickly during the late 1980s from an environment characterized by trust and confidence in the inherent quality of medical care to an environment characterized by the presumption of variation in quality and the demand for hard evidence on quality from health plans and providers.

### **C. Unique Significance of Networks for Chronic Care Populations**

These developments have had a particular significance for patients with chronic illness or need for chronic care, and their employers and communities. Chronic illness typically implies not only more frequent and more intensive provision of healthcare, but a higher likelihood of need for a mix of services across the care continuum and the need for contact with more individual providers and more physical locations for service.<sup>15,16</sup> Organized delivery systems offer the theoretical advantage of “seamless,” more highly coordinated care across the continuum and across a diverse set of providers and care sites. An organization that can provide inpatient, outpatient, home health, rehabilitation, and other forms of care under the same administrative structure, using the same medical records system, with a single patient registration and billing system, and a coherent approach to practice patterns and guidelines, presumably offers patients with chronic illness a more coordinated, rational, effective pattern of care. There is still a lack of compelling evidence that these theoretical advantages regularly occur in practice, but “success stories” are relatively easy to find.<sup>17,18</sup>

As patients and families make choices among competing networks, plans, or other kinds of organized delivery systems (inevitable in many cases given employers’ preferences for managed care and/or the affiliation of providers into organized systems in a given area), they face many of the same information needs as do large purchasers. Who offers the best balance of cost and quality? Whose patients have the best clinical outcomes? Whose patients are most satisfied with the care they receive? Whose services are most accessible and convenient? Which network or system includes the physician(s) I want to see?

### **D. Health Care “Report Cards”**

In response to these needs for information to make informed choices, a number of “report card” or “performance report” initiatives have begun. The fundamental purpose of these initiatives is to provide a set of quantitative, meaningful, comparable indicators of quality that will allow purchasers and individual consumers to make informed decisions. The Consumer Reports model of scores and rankings for competing products has been a template for many of these reports. The underlying premise is that quality can be measured and compared across organizations on a number of dimensions, and that these comparisons will allow better, more informed choices by consumers.

United Healthcare<sup>19</sup> and Kaiser<sup>20</sup> were among the first health plans or delivery organizations to publish a report card on their own performance. Several business coalitions in major cities began compiling and reporting comparative performance data (e.g., Cleveland Health Quality Choice project<sup>21</sup> and Buyers Health Care Action Group in the Twin Cities<sup>22</sup>). Healthcare consulting firms began to define quality measures and include requests for this information from plans and providers as part of their role in advising corporate benefits managers.

Very quickly, again in the late 1980s or very early 1990s, this activity created a “Tower of Babel” situation in which measurement systems of all shapes and sizes were being offered or recommended to purchasers, and plans and providers were being asked for information by purchasers that was either redundant, only slightly different from request to request, or totally different for each purchaser/consultant. It was obvious that there was a need for some sort of standardization in the area of quality measures and report cards, based on some general agreement among the users of the information about what really mattered.<sup>23</sup>

## E. Performance Measurement Systems

Several prominent quality measurement systems came about as a result of this desire for standardization. The best known is the Health Plan Employer Data and Information Set (HEDIS), which was originally developed by the HMO Group and a group of large national employers in 1991, then adopted, refined, and expanded by the National Committee for Quality Assurance (NCQA).<sup>24,25</sup> This measure set, which is now in its fourth significant iteration, is designed to measure quality of health plans for use by purchasers and individual consumers (see detailed description below). The Joint Commission for Accreditation of Healthcare Organizations (JCAHO) developed an extensive set of quality measures, mainly for use by hospitals, under the name “IMSystem.” This measure set was extensively tested but never fully implemented because of hospitals’ concerns about burden of data collection and utility of the information. A remnant of the system still exists under the label “ORYX Plus” as an option in JCAHO’s new system in which hospitals can choose from among a long list of approved measurement systems.<sup>26</sup>

In 1994, many of the same large private purchasers who had helped develop HEDIS met with Paul Ellwood to establish the Foundation for Accountability (FACCT). The motivational force behind the group’s formation was the observation that HEDIS and other similar measurement systems did not include much information on health outcomes, and therefore fell short of the vision of accountability for outcomes that Dr. Ellwood described in his 1988 *New England Journal of Medicine* article.<sup>12</sup> FACCT commissioned the development of measure sets of several different disease areas with an emphasis on health status outcomes (see detailed description below).<sup>27</sup>

In 1991, researchers at Henry Ford Health System in Detroit published a monograph through the American Hospital Association describing a philosophy of measurement and a set of 91 specific performance measures for vertically integrated healthcare systems.<sup>28</sup> In 1992, several integrated systems from around the country formed the Consortium for Research on Indicators of System Performance (CRISP) to test, refine, and use many of these measures in the context of benchmarking and internal quality improvement.<sup>29,30</sup> The CRISP measure set is continually being revised through the experience of the consortium, but represents another model for measuring and comparing performance of organized delivery systems.

At about the same time, Steven Shortell, his colleagues at Northwestern University, and collaborators at KPMG/Peat Marwick organized the Health System Integration Study (HSIS) with the active participation of nine large vertically integrated health systems. The focus of HSIS was not on performance measurement per se, but on the internal dynamics of integrated systems, the degree to which organizations had achieved “systemness” in several domains, and the relationship between measures of “systemness” and selected measures of financial and clinical performance.<sup>31,32</sup> The study produced two sets of measures that can appropriately be called performance measures, one based on ratings of key system “stakeholders” about the system’s internal organization and function, and one based on more objective data on financial and clinical performance.

The National Chronic Care Consortium (NCCC), formed in 1991, took up a similar effort as one of its first major projects. The group developed the Self-Assessment for Systems Integration (SASI™) tool as a way for member organizations (and others) to examine the extent to which they had created, and were functioning, as a truly integrated system of care

for the chronically ill. The SASI tool is an inventory of capabilities, relationships, and dimensions of performance with a specific focus on chronic care.<sup>33</sup> The SASI tool is similar to the HSIS set of measures in that it focuses on the healthcare organization itself and the relationships among its parts rather than focusing on quantitative measures of healthcare processes and outcomes in a denominator population of patients or members.

All of these initiatives, and all of the trends discussed thus far, occurred in an environment where industrial concepts of Total Quality Management (TQM) or Continuous Quality Improvement (CQI) were beginning to dominate the healthcare field.<sup>34,35</sup> CQI's emphasis on healthcare as a process, with formal measurement of key process characteristics, created both an increased capacity for measuring and reporting performance, and an emphasis on comparative use of data for "benchmarking" and identification of "best practices."<sup>36,37</sup> CQI principles were adopted during the 1980s first in manufacturing settings, then in service industries, then in healthcare under a feeling of intense competitive pressure and the need to improve quality as a means of organizational survival.<sup>38</sup> Organizations wishing to track their progress toward a CQI culture were able to use a set of standards used by the U.S. Commerce Department for judging applicants for its Malcolm Baldrige National Quality Awards.<sup>39</sup> These standards or criteria are similar to the HSIS or SASI measures in that they involve a close examination of organizational capabilities, internal relationships among departments, and daily practices.

## **F. Relative Strengths and Weaknesses**

There are, then, at least six prominent, widely used measurement systems that can be used to compare healthcare plans or networks. These measurement systems were all developed at approximately the same time (late 1980s to early 1990s), but for somewhat different purposes and with different primary users in mind. For organizations considering investment in one or more of these systems, which one represents the best choice? For what purpose(s)? With what benefit(s), given the cost of data collection and analysis? The purpose of this monograph is to describe these measurement systems in detail, indicate similarities and differences between them, make suggestions about appropriate uses of the different systems, and suggest directions for the further development of the performance measurement field in general.

## II. Conceptual Foundation

### A. Chronic Care Patients—Key Dimensions of Performance

No judgment about the relative merits of competing measurement systems, or choice among the systems for implementation and use, can be made without some sort of conceptual foundation. For chronic care patients, the unique needs and requirements of the healthcare systems provide the basis for such a foundation. These unique needs include:

- continuity and coordination of care across multiple specialties, individual providers, and sites of care;
- the management of functional health status, independence, and ability to perform normal daily activities in addition to management of the biological aspects of disease;
- coordination of supportive social and psychological services with medical services;
- frequent monitoring of disease status (e.g., blood glucose monitoring for diabetes, prothrombin time measurement for patients with clotting disorders);
- planning for predictable declines in health in later stages of the disease course;
- prevention of acute exacerbations of disease;
- provision of “sub-acute” residential services for purposes of either caregiver respite, stabilization of disease, or humane management of terminal care;
- patient/family experience of illness as long, continuous episodes rather than isolated clinical encounters;
- the need to manage information, both from provider to patient and from provider to provider.

If these features of patient needs or organizational priority are a reasonable description of the “mission” of chronic care networks, then a measurement system must include the ability to assess:

- health, function, and independence, not necessarily cure of disease;
- coordination among multiple providers and sites of care;
- prevention of acute exacerbations;
- costs for whole episode or period of time, not individual unit of service;
- satisfaction;
- quality of care vis-a-vis accepted guidelines and standards, including the provision of necessary services and the avoidance of unnecessary or futile services;
- extent to which care is provided with a minimum of “administrative hassles”;
- the overall “burden of illness” for patients, families, and other caregivers.

## **B. Networks**

The selection of measures cannot be divorced from some understanding about the organization being measured. Not all of the measure sets are equally appropriate for all organizational forms, and some are specifically designed for a single form (e.g., HEDIS for health plans, HSIS for vertically integrated health systems). Although none of the organizational forms will be excluded from discussion in the rest of this monograph, the discussion of measurement systems will have in mind a “prototype” chronic care network with the following features:

### **1. Emphasis on Delivery of Care, Not Insurance**

Even if the chronic care network is, by title, license, and function a health plan, the discussion of measures will emphasize the process and outcome of healthcare rather than the health plan’s insurance and member relations functions.

### **2. Centralized Budgeting, Quality Improvement, and Information Systems**

The discussion of measures will presume that the organization has the ability to use some centralized administrative or clinical databases to measure the care provided to chronic care patients and to act on the information through financial and management control mechanisms to influence the care process.

### **3. Broad Accountability for Cost, Outcomes, Quality, and Satisfaction**

The organization does not simply provide one specific clinical service (e.g., home healthcare), but rather a comprehensive spectrum of services that implies a comprehensive responsibility for patient outcomes.

### **4. Financing Arrangements**

The network may be paid through a variety of arrangements that can include capitation, fee-for-service, or variants of these basic models. Regardless of specific payment type, the network has some tangible incentive for cost containment and efficiency, constrained by the obligation to provide high-quality care and maintain high levels of patient satisfaction.

## **C. Desirable Features of a Performance Measurement System**

With these general concepts about chronic care patients and chronic care networks in mind, the following general points can be made about measurement systems that can guide evaluation:

### **1. Systems Should Be Designed Around Clear Purpose and Target Audience**

Not only must the measurement system have some clear purpose and target audience, that purpose and target audience must be reasonably well-aligned with chronic care. The APACHE III system for measuring ICU care relative to predicted patient need and outcome, for example,<sup>40</sup> is an excellent system for its defined purpose, but it is not closely enough aligned to the typical experience of chronic care patients to be included in this discussion. However, there is still room for acceptable variation in intended use for a measurement system for chronic care networks. Some major options include:

- *Internal Improvement or Benchmarking.* Measures designed for this purpose are typically clinically specific, tightly linked to a defined process or related set of processes, collected as close to “real time” as possible, responsive to changes made in the care process itself, and expressed as run charts or some similar method that highlights trends over frequent time intervals.

- *Consumer Choice.* Measures designed for this purpose are less clinically specific, although there may be value in disease-specific measures for patients with chronic illness and the future expectation of frequent care for that illness (i.e., patients with diabetes will be interested in quality of care for diabetes rather than quality of care in general). Patients are interested in how others similar to them (e.g., other diabetic patients) judge quality of care, so there is greater use of patient satisfaction and patient reports on quality of care than on highly technical quality measures.<sup>41,42</sup>

Corporate purchasers have similar interests to the extent that they are acting as agent for a collection of individual employees, but they are more interested in “population” or “group” measures that reflect efficiency and value at a more aggregate level (e.g., rates of hospitalization, rates of return to work following surgery).<sup>43</sup>

For both individual and corporate consumers, it is essential that measures reflect stable organizational features that distinguish organizations in the same way, with the same relative rankings, over relatively long periods of time (e.g., three years). If this condition does not hold, decisions made this year based on last year’s data have no predictive accuracy for next year’s experience.

- *Accreditation.* Measures used to meet accreditation requirements (e.g., NCQA, JCAHO) are similar to those used for corporate purchasers in that they should be relatively stable and potentially different among organizations being ranked or rated. Since accreditation is typically a “pass/fail” system, though, measures designed for accreditation can be designed to distinguish only among two or three levels of quality rather than to rank organizations that all have “acceptable” levels of quality.

## **2. Measures Are Defined in Quantitative, Objective Terms**

A collection of verbal statements, anecdotes, testimonials, or case examples can be used to paint a very compelling picture of performance, either high or low. However, this approach in a performance measurement system makes the comparisons dependent on the organizations’ ability to find, collect, and “package” the information rather than on performance per se. Measures that are clearly and objectively defined and expressed in quantitative terms are far less dependent on the organizations’ skill in “information packaging.”

## **3. Measures, Taken Together, Provide a Comprehensive “Picture” of the Network**

A set of measures focusing on only one or two aspects of network performance is unlikely to be informative for consumer choice or accreditation purposes. Such an approach may be appropriate for internal improvement purposes, although over time it would be expected that a more comprehensive system would be needed to support attention to other areas as well. Criticism of HEDIS for its focus on process of care and preventive services measures was one of the reasons for the formation of FACCT.<sup>44</sup>

## **4. There is Documented Reliability, Validity, and Sensitivity**

The measures must ultimately “tell the truth.” That is, they must not be subject to too much random error, be biased, be reflective of something other than what they claim, or lag too far behind meaningful changes in organization or process of care.

## **5. Adjustments Are Made for Case Mix or Disease Severity**

In Garrison Keillor’s Lake Wobegon, all the children are above average. In healthcare, everyone’s patients are sicker than average. Performance, particularly in the area of patient outcomes, clearly depends on the underlying severity or complexity of illness in the

population served. However, since everyone's patients cannot be sicker than average, the measures must have some mechanism for adjustment on the basis of case mix. This technology is relatively well-developed for hospital care,<sup>45</sup> but is not as far along for assessing ambulatory and long-term care in chronically ill populations.<sup>46,47</sup> Measurement systems that include case-mix or severity adjustment of proven validity are preferable to those that do not, other considerations being equal.

## 6. Appeal to Users

Regardless of the intended audience—patients, purchasers, clinicians, administrators—a measure set has to have some intuitive appeal and be relatively easy to use. More specifically, this means:

- *“Face Validity” Basis in Peer Experience for Lay Users.* The meaning and underlying purpose of a measure should be clear (ideally obvious) to the users. Patients have a very clear understanding about measures of satisfaction or wait time for appointments, but have much less “feel” for measures of mean glycosylated hemoglobin level or hospital readmission rate.<sup>48</sup> Clinicians, on the other hand, may have the opposite preferences because of their higher level of technical knowledge and their focus on clinical parameters.
- *Clinical Detail for Acceptance by Clinicians.* Clinicians are also particularly sensitive to the importance of different severity levels, comorbidities, demographic and social factors, and possible genetic or environmental factors in evaluating processes and outcomes of care.<sup>49,50</sup> A measure of glycemic control for all diabetics in all settings, for example, is less likely to be accepted and used than a measure of glycemic control for Type I diabetics stratified by age, duration of disease, and site of care.
- *Stability over Time and Link to Key “Decision Dimensions” for Purchasers.* Purchasers making decisions about which health plans to offer employees or which networks to include in “carve-out” contracts require that the information included in a measure set be timely, predictive of employees’ future experience, and reflective of experience of a broad spectrum of employees affected. For example, measures of excellence of ICU care may not be as useful as more global measures of patient satisfaction with access, since most employees will seek access to care of some kind in a period of time, but very few will be admitted to the ICU.

---

## III. Description of Network Performance Measurement Systems

### A. Self-Assessment for Systems Integration (SASI™)

#### 1. History and Purpose

The SASI tool was created in 1995 by a task force representing member organizations of the NCCC. The SASI User's Guide defines it as "a tool/resource for health systems and other health provider organizations to use as they work toward integrating the components of the system in order to better serve people with chronic conditions . . . the tool is conceived as one (original emphasis) resource that the organization can use to plan and evaluate its activities toward network integration in serving people with chronic conditions."

The SASI tool was designed to meet a need for member healthcare systems of the NCCC and other emerging healthcare networks to take stock of their current level of system integration and identify opportunities for future improvement. It is based on the premise that higher levels of integration are likely to result in better care for chronically ill patients, and it provides a vehicle for assessing levels of integration in several domains. Its primary audience is internal—that is, the network itself gathers, analyzes, and uses the information. The format of the tool and the measure set allows for "benchmarking" with other chronic care networks, but the SASI tool could be used by a single network comparing its current standard of system integration to some ideal or desired state. The SASI tool includes a "Resource Guide" that allows networks to plan and implement improvements identified as priorities in the measurement parts of the tool.

#### 2. General Structure

There are four major sections to SASI. There is a definition section identifying key performance domains and indicators within those domains, a set of "templates" or "forms" in which network performance can be assessed in terms of a large number of indicators of system integration, and a resource guide to support planning and implementation of improvements.

The assessments could conceivably be done by a single individual, either an employee of the organization or an external consultant, but the tool was designed with a team in mind—a team of clinicians and administrators working together as part of a strategic planning, quality improvement, or ongoing assessment initiative. These individuals would share perspectives and objective information to arrive at some sort of consensus about the state of network integration in each indicator.

#### 3. Main Sections or Important Subdivisions

Section I: *Guidelines and Indicators* defines the key features and performance domains for integrated chronic care networks. This section is organized around three major "elements of integration" (e.g., network infrastructure), which are subdivided into nine "objectives" (e.g., to provide a full array of effective and efficient services). The objectives are broad and are used to set targets for measurement in more detail.

Figure 1

## Introductory Notes to the SASI™ Tool

The SASI™ tool provides a framework for conceptualizing, implementing, and measuring chronic care integration strategies used by a network of health, long-term care, and community providers:

- **Chronic care integration strategies** are those methods and processes used to improve the ability of individual care providers to work together as a single system in serving persons with chronic illness.
- A “**chronic care network**” (CCN) is a set of providers working together to implement chronic care integration strategies.
- A CCN can take many forms. It could, for example, be a subset of a larger healthcare network, or it could be an entire healthcare network that has instituted all CCN components throughout its system.
- A **healthcare network** is broadly defined to be a strategic alliance between an array of health and long-term care providers and community support services.
- A **client** is broadly defined as an individual having a chronic condition **and** her or his caregiver(s).

In the NCCC’s vision of a CCN, the CCN’s work is centered around nine chronic care integration strategies. This tool outlines these nine strategies that support a CCN regardless of the CCN’s specific structural form. SASI addresses three core issues:

- establishing systems and processes that promote provider collaboration across sites toward common goals;
- providing the full continuum of chronic care services; and
- managing service use to assure quality and cost-effectiveness.

The primary audience for this tool is management and leadership staff who are responsible for improving the integration of chronic care. Certain sections may be especially relevant to clinical teams comprised of primary care physicians, nurses, therapists, and social work staff. Other sections are more relevant to administrative staff responsible for other functions: clinic administrators or directors of finance, risk contracting, information systems, customer service, marketing, and strategic planning. Interdisciplinary and cross-site teams are important and are referenced frequently.

Figure 2

<h2 style="text-align: center;">Overview of SASI™ Essential Elements and Objectives</h2>		
Elements of Integration	Description	Objectives
CCN Infrastructure	CCN systems and processes promote provider collaboration across sites toward common goals.	<ol style="list-style-type: none"> <li>1. <b>Governance structures</b> support goal development and improve the ability of individual care providers to work together as a single system.</li> <li>2. <b>Management strategies and structures</b> support cross-site, interdisciplinary integration efforts.</li> <li>3. <b>Information technology systems</b> allow providers in all settings to share meaningful information about clients, costs, and operations.</li> <li>4. <b>Financing systems</b> promote systemwide management of cumulative costs, tied to care outcomes.</li> </ol>
A Full Array of Services	The CCN provides a full continuum of chronic care services.	<ol style="list-style-type: none"> <li>5. The needs of <b>high-risk populations</b> are identified.</li> <li>6. <b>A full array</b> of effective and efficient services is provided.</li> </ol>
Service Allocation Strategy	The CCN manages service use to assure the quality, appropriateness, and cost-effectiveness of care.	<ol style="list-style-type: none"> <li>7. Care management is focused on <b>disability prevention</b> and organized around <b>defined populations</b> (e.g., high-risk, condition-specific).</li> <li>8. <b>Seamless care</b> is provided across settings and over time.</li> <li>9. <b>Clients are involved</b> in care management and self-care activities.</li> </ol>

Section II: *Workbook* contains the detailed assessment tool for measuring progress in specific activities related to the nine objectives. The specific activities for each objective have to do with either conceptualizing, implementing, or measuring progress toward that objective. In each specific performance area listed, there is space on a worksheet for assessing “What progress has our network made in this area?”, “What areas need to be improved?”, and “What steps will our network take to improve this indicator in the next year?”. Examples of indicators include:

- practitioners and staff throughout the network participate in training regarding the meaning and implementation of integrated chronic care;
- medical records have a unique identifier for members across settings;
- adverse outcomes for the network are tracked to assess the need for improvements in condition-specific care.

No quantitative rating scale or scoring system is provided for these indicators. The open-ended format would allow the use of a rating scale or scoring system if an organization wished to develop one, but the format also allows use of text, examples, verbal descriptors, or essentially any other way of describing the current status of the network and future plans on each indicator.

Section III: *Global Measures* includes a structured survey with a mix of multiple-choice or Likert-type rating scales to assess many different aspects of network capability and level of integration. These questions are linked closely to the three elements of integration and nine objectives in the prior two sections. Examples of survey questions include:

- Prevention services are adequately funded. (Four-step rating scale from “very true” to “not generally true.”)
- Does the network have a system for intake and registration that is shared across settings?
- What percent of clients with serious and persistent chronic conditions have their care managed through condition-based care management methods? (seven-step rating scale covering a response range of zero percent to 100 percent.)

Section IV: *Resource Guide* is not part of the measurement system per se, but provides a list of resources (printed materials, phone numbers, groups to contact) organized around the nine objectives that run as consistent themes throughout SASI. This part of the SASI tool is designed to support improvement efforts that are identified through the prior assessment steps.

#### **4. Specific Measures**

See Figure 3: Sample Questions from the Global Measures Section of the SASI™ Tool

#### **5. Current status**

The SASI tool was formally tested in early 1995 by a set of volunteer organizations drawn from the NCCC membership. It is available for use by NCCC members and other interested organizations. The SASI tool has been used by at least 11 NCCC member organizations as of the fall of 1997, and four other member organizations have used at least part of it for formal self-assessment. The specific types of self-assessment activities have varied. One member organization used SASI as the framework for a one-day retreat on improving diabetes care for residents of the surrounding community, while another used SASI as a framework for identifying strengths and weaknesses among potential merger partners.

**Figure 3**

**Sample Questions from the  
Global Measures Section of the SASI™ Tool**

- Does the healthcare network have a CCN strategic plan?
- Does the network regularly measure progress toward chronic care integration?
- Does the network have specific CCN work plans?
- What percent of service settings within the network are under a common accounting, strategic planning, budgeting, and fiscal reporting cycle?
- Does the network have a core set of care and cost data that are collected across settings?
- Does the network have specific tools for identifying high-risk populations?
- Does the network provide any special training to practitioners and teams about transitional issues?
- Has the network documented the effects of education, self-help, self-care services on client satisfaction, care outcomes, self-care practices, and/or costs in any way over the last two years?

## B. NCQA/Health Plan Employer Data and Information Set (HEDIS )

### 1. History and Purpose

HEDIS was originally developed by a group of large staff- and group-model HMOs (the HMO Group) and a group of large private firms (Bull HN Information Systems, Digital Equipment Corp., Xerox, and GTE) interested in standardizing the data requirements for health plans. The stated objectives were to: “(1) define and understand employer needs to document the “value” of a health plan; and (2) develop performance measures that would provide data and information in response to those needs.”<sup>51</sup>

The original development effort began in 1989, and a draft document known as HEDIS 1.0 was released in 1991. Shortly thereafter, the NCQA took over responsibility for HEDIS and began a series of revisions, expansions, and refinements in response to feedback from users. A Performance Assessment Committee (PAC) was established to oversee this effort. The PAC was made up of clinicians and quality experts from large health plans and health benefits managers from the four companies that were the original sponsors of the effort. This group began work on HEDIS 2.0 in late 1992 and released that version of HEDIS in 1993. A national pilot study of HEDIS 2.0 was conducted in 1994 in approximately 20 volunteer health plans, and some revisions were made to the technical specifications of the measures, resulting in a HEDIS 2.5 released in 1995. A version of HEDIS specific to Medicaid populations was developed, and the NCQA supported groups of researchers in the development of expanded measures for several specific chronic diseases. (This effort is still underway.)

HEDIS 2.5 very rapidly became the standard for performance measures of health plans, and by 1996 there was no serious competing system. Most corporate purchasers or purchasing coalitions required health plans to provide HEDIS data if they had a requirement for comparative performance data in their requests for proposals. Although not all plans could report all HEDIS measures (very small plans and some IPA-model plans do not necessarily have the required data capabilities), HEDIS 2.5 became, and still is, the standard performance measure set for health plans.

In 1995, the NCQA established a Committee on Performance Measurement (CPM) to carry out the expansion and revision functions that had been carried out by the PAC during development of HEDIS 2.0. This group developed, and released for public comment, an expanded measure set called HEDIS 3.0 in July of 1996. HEDIS 3.0 includes more measures of outcome than the prior versions, and includes the concept of a “Testing Set” of measures that are recommended for trial use but not required. HEDIS 3.0 has not yet been fully accepted as the standard for health plan measurement because of some concerns about the additional burden of data collection compared to HEDIS 2.5, but it is the current standard in this area by virtue of being the most recent version of the measure set released with full NCQA support.

### 2. General Structure

HEDIS 3.0 (and prior versions) is organized around eight major domains of health plan performance. These are:

1. *Effectiveness of Care*. This domain includes several specific measures of services provided in accordance with accepted guidelines for specific clinical conditions (e.g., mammography in women ages 52 to 69). The measures are not effectiveness measures per se (i.e., they do not measure patient outcomes), but they reflect the provision of services generally acknowledged to be effective.
2. *Access to Care*. This domain includes measures of access to appointments with primary care physicians and specialists, access to telephone information, and early initiation of prenatal care.

3. *Satisfaction with the Experience of Care.* Results of an annual, standardized member satisfaction survey.
4. *Health Plan Stability.* This domain includes measures of member and physician turnover as well as information on financial performance and rate trends.
5. *Use of Services.* Several measures of utilization (admissions, visits, specialty referrals, procedures performed) for specific clinical conditions. In contrast to the “Effectiveness” section, many of these events are “undesirable” (e.g., readmission for specified mental health disorders) so that low rather than high rates are required.
6. *Cost of Care.* Rate trends and rates of occurrence of high-cost DRGs.
7. *Informed Health Care Choices.* This domain includes measures of information available to members, including foreign language translation services.
8. *Health Plan Descriptive Information.* A blend of quantitative and nonquantitative information on member services, credentialing and quality management capabilities, characteristics of plan members, and range of benefits/services provided.

Within each domain, a set of specific measures or indicators is listed. The measures are typically expressed as rates (e.g., cesarean section or VBAC rates), with data drawn from either administrative databases (e.g., claims) or medical records. In the domain of health plan descriptive information, the information is more text-oriented or yes/no format rather than quantitative.

In all domains, the health plan is the organizational unit being measured. HEDIS measures can be used to evaluate medical groups, provider networks, regions of a health plan, or other subdivisions of a plan. The measures are designed, though, to measure health plans with relatively large, identifiable denominator populations that are required to calculate rates of performance in several specific disease domains.

The measures reflect care of both adults and children, and emphasize the provision of preventive services. This emphasis reflects a concern on the part of employers/purchasers that managed care incentives (particularly capitation payment) may lead to poor access or underservice of enrolled populations. The measures are designed to identify instances of poor access to care, poor provision of preventive services, low patient/member satisfaction, or overuse of expensive services (e.g., hospitalization) that may reflect lack of access to less expensive levels of care (e.g., primary care physician).

### 3. Main Sections or Important Subdivisions

HEDIS 3.0 includes a “Reporting Set” of measures and a “Testing Set.” The reporting set is the list of “required” measures—those that have detailed specifications and sufficient track records to be considered “finished” and appropriate as part of a standard package of information shared between health plans and employers. The testing set is a set of measures that have unresolved questions in terms of interpretation of data, definition of denominator populations, need for case-mix adjustment, or appropriateness in comparing health plans. The NCQA recommends that health plans collect the testing set measures if possible, and use the experience to further refine the measures, understand their value, and eventually move as many of them as possible to the reporting set. Both the testing set and the reporting set include measures from the eight major performance domains listed above.

### 4. Specific Measures

See Figure 4: Health Plan Employer Data and Information Set (HEDIS) 3.0 Reporting Set Measures

Figure 4

## Health Plan Employer Data and Information Set (HEDIS) 3.0 Reporting Set Measures

<b>Domain: Effectiveness of Care</b>	
Childhood Immunization Status Adolescent Immunization Status Advising Smokers to Quit Flu Shots for Older Adults Breast Cancer Screening Cervical Cancer Screening Prenatal Care in the First Trimester	Check-Ups After Delivery Treating Children's Ear Infections Beta Blocker Treatment After a Heart Attack Eye Exams for People with Diabetes The Health of Seniors Follow-Up After Hospitalization for Mental Illness Low Birth-Weight Babies*
<b>Domain: Access/Availability of Care</b>	
Adults' Access to Preventive/Ambulatory Health Services Children's Access to Primary Care Providers Availability of Primary Care Providers Availability of Mental Health/Chemical Dependency Providers Availability of Obstetrical and Prenatal Care Providers	Initiation of Prenatal Care Low Birth-Weight Deliveries at Facilities for High-Risk Deliveries and Neonates Annual Dental Visit Availability of Dentists Availability of Language Interpretation Services
<b>Domain: Satisfaction with the Experience of Care</b>	
Member Satisfaction Survey	Survey Descriptive Information
<b>Domain: Health Plan Stability</b>	
Disenrollment Provider Turnover Years in Business/Total Membership	Indicators of Financial Stability Narrative Information on Rate Trends, Financial Stability, and Insolvency Protection
<b>Domain: Use Of Services</b>	
Frequency of Ongoing Prenatal Care Well-Child Visits in the First 15 Months of Life Well-Child Visits in the Third, Fourth, Fifth, and Sixth Year of Life Adolescent Well-Care Visit Frequency of Selected Procedures Inpatient Utilization—General Hospital/Acute Care Ambulatory Care Inpatient Utilization—Non-acute Care Discharge and Average Length of Stay—Maternity Care Cesarean Section and Vaginal Birth After Cesarean Rate (VBAC-Rate)* Births and Average Length of Stay, Newborns	Mental Health Utilization—Inpatient Design and Average Length of Stay Mental Health Utilization—Percentage of Members Receiving Inpatient, Day/Night, and Ambulatory Services Readmission for Specified Mental Health Disorders Chemical Dependency Utilization—Inpatient Discharges and Average Length of Stay Chemical Dependency Utilization—Percentage of Members Receiving Inpatient, Day/Night, and Ambulatory Services Readmission for Chemical Dependency Outpatient Drug Utilization
<b>Domain: Cost of Care</b>	
Rate Trends	High-Occurrence/High-Cost DRGs
<b>Domain: Informed Health Care Choices</b>	
New Member Orientation/Education	Language Translation Services

\*The low birth-weight measures and VBAC rate will remain in the Reporting Set, but plans will not be expected to report them for Reporting Year 1996.

## HEDIS 3.0 Reporting Set Measures

### Domain: Health Plan Descriptive Information

Board Certification/Residency Completion	Arrangements with Public Health, Educational, and Social Service Organizations
Provider Compensation	Pediatric Mental Health Network
Physicians Under Capitation	Chemical Dependency Services
Case Management	Family Planning Services
Utilization Management	Total Enrollment
Risk Management	Enrollment by Payer (Member Years/Months)
Quality Assessment and Improvement	Unduplicated Count of Medicaid Members
Recredentialing	Cultural Diversity of Medicaid Membership
Preventive Care and Health Promotion	Weeks of Pregnancy at Time of Enrollment in the Health Plan

## HEDIS 3.0 Testing Set Measures

### Domain: Effectiveness of Care

Substance Counseling for Adolescents	Controlling High Blood Pressure
Number of People in the Plan Who Smoke	Assessment of How Breast Cancer Therapy Affects the Patient's Ability to Function
Smokers Who Quit	Prescription of Antibiotics for the Prevention of HIV-Related Pneumonia
Flu Shots for High-Risk Adults	Screening for Chemical Dependency
Stage at Which Breast Cancer Was Detected	Continuity of Care for Substance Abuse Patients
Chlamydia Screening	Failure of Substance Abuse Treatment
Colorectal Cancer Screening	Continuation of Depression Treatment
Aspirin Treatment After a Heart Attack	Availability of Medication Management and Psychotherapy for Patients with Schizophrenia
Follow-Up After an Abnormal Pap Smear	Appropriate Use of Psychotherapeutic Medications
FollowUp After Abnormal Mammogram	Family Visit for Children 12 Years of Age or Younger
Use of Appropriate Medications for People with Asthma	Patient Satisfaction with Mental Health Care
Monitoring Diabetes Patients	
Prevention of Stroke in People with Atrial Fibrillation	
Outpatient Care of Patients Hospitalized for Heart Failure	
Cholesterol Management of Patients Hospitalized for Coronary Artery Disease	

### Domain: Access/Availability of Care

Problems with Obtaining Care

### Domain: Satisfaction with the Experience Of Care

Consumer Assessments of Health Plan Study (CAHPS)	Disenrollment Survey Satisfaction with Breast Cancer Treatment
---	---

### Domain: Use of Services

Use of Behavioral Health Services

### Domain: Cost of Care

Health Plan Costs per Member per Month

### Domain: Informed Health Care Choices

Counseling Women About Hormone Replacement Therapy

Source: *HEDIS 3.0*. Washington, DC: National Committee for Quality Assurance, January 1997, pp 15-18: Reprinted with permission.

## 5. Current Status

HEDIS 2.5 has been used by hundreds of health plans and their associated provider networks, and most are making a transition to HEDIS 3.0. There are a number of formal HEDIS “benchmarking” activities in states or local market areas.<sup>52</sup> Some large national employers (e.g., General Motors) are beginning to incorporate HEDIS measures in either their choices of health plans to offer or the size of the monthly contribution to employee health plan premiums. The goal of these latter efforts is to steer employees to “benchmark” or “high-quality” plans rather than plans with lower performance ratings. HEDIS measures are also beginning to be used to compare provider networks within health plans<sup>53</sup> or to compare group practices in areas where providers typically contract with all or most health plans in an area.<sup>54</sup>

The Committee on Performance Measurement continues to meet regularly to discuss refinements to HEDIS and to take up questions about use of HEDIS for specific clinical or insured populations (e.g., Medicare Risk contracts, Medicaid, patients with chronic illness).

## C. Health System Integration Study

### 1. History and Purpose

The Health System Integration Study (HSIS) was conducted between 1990 and 1993 by a team of investigators from Northwestern University (Steven Shortell, Ph.D., Principal Investigator) and KPMG/Peat Marwick.<sup>55</sup> The study was done on, and in collaboration with, nine large, vertically integrated healthcare systems:

- Baylor Health Care System, Dallas, TX
- Evangelical Health Systems, Oakbrook, IL
- Fairview Hospital and Healthcare Services, Minneapolis, MN
- Franciscan Health System, Aston, PA
- Henry Ford Health System, Detroit, MI
- Sharp HealthCare, San Diego, CA
- Sisters of Providence, Seattle, WA
- Sutter Health, Sacramento, CA
- UniHealth America, Burbank, CA

The purpose of the study was to understand “systemness” in several specific domains (e.g., integration of information systems), to develop measures of “systemness,” and to study the relationship between measures of “systemness” and other measures of organizational performance (e.g., financial performance or quality of care).

The study involved detailed background surveys about the history, mission, organizational structure, and governance structure of the participating systems (late 1990), detailed surveys regarding perceptions of “systemness” completed by physician leaders, administrators, and members of boards of trustees (1991), site visits by project investigators to interview many survey respondents and review relevant documents (1991), and follow-up visits to collect data on “objective” performance measures (1992-1993). Results have been reported in published articles and more detailed technical reports since 1994.<sup>56,57</sup>

### 2. General Structure

There are two major sets of measures in the HSIS. The survey measures of system integration are made up of a large number of Likert-type rating scales that reflect various aspects of integration (e.g., extent to which clinical guidelines are disseminated throughout the system). The surveys are designed to be completed by individuals with different perspectives or positions in the health system—clinicians, administrators, board members, etc. By asking samples of individuals in the different respondent groups to complete the same surveys, it is possible to compare perceptions of integration across the groups. It is possible, for example,

to determine whether clinicians and administrators differ in their views of the extent to which the system has achieved high levels of clinical integration.

The set of “objective” performance measures includes a detailed set of financial, clinical quality (e.g., hospital-acquired infection rates), and utilization (e.g., percent of support services shared among system operating units) measures. These measures are calculated from a variety of data sources.

The measures are designed to reflect integration or performance at the system level. Most are not suitable or meaningful for evaluation of smaller organizational units—departments, divisions, or regional subsystems. The measures generally do not require the identification of specific denominator populations of patients or plan members, but some measures require the identification of a population of system-affiliated physicians for whom certain services are either provided or not. The measures were not designed for a specific focus on chronic illness or chronic care populations, but many of the system integration concepts and dimensions are highly relevant to that patient population.

### 3. Main Sections or Important Subdivisions

As indicated above, there are two major sets of measures: those having to do with perceptions of integration and those having to do with objective performance. In the first major set, there are the following subdivisions (survey sections or separate surveys):

- *Strategy Measures.* This survey includes approximately 70 questions having to do with respondents’ perceptions of the systems’ strategic plans and priorities among possible strategic alternatives.
- *Perceived Systemness and Integration Measures.* This survey includes an extensive set of questions about perceived integration among system entities in areas of clinical services, strategic planning, information systems, quality assurance, staff development, and others. Survey responses are scored in a way that allows “perceived integration” scores to be reported separately for each specific domain of system integration.
- *System Office–Unit Relationship Measures.* This survey includes 26 items having to do with the roles and relationships of “central office” and “field” units in areas like capital allocation, clinical service development, staff development and selection, etc.

In the objectives measures set, there is a similar categorization of measures into categories that reflect major concepts or domains of system integration. These include:

- *Functional Integration.* These measures reflect the extent to which support functions and activities (e.g., billing, human resources) are coordinated or shared across business units within a system. Data came from a “yes/no” survey completed by system strategic planners.
- *Physician-System Integration.* This measure set includes measures of the extent to which physicians focus their activities within the system or at system-affiliated facilities, and the extent to which the system provides support services for physicians. The set is further divided into four subcomponents: economic involvement (e.g., admissions by physicians to system-affiliated hospitals, scope and extent of joint venture activity); administrative involvement (e.g., physician presence on governing boards); group practice formation (e.g., percent of physicians in group practice); and shared accountability (e.g., percent of system units with a single, common medical staff organization).
- *Clinical Integration.* This set of measures reflects the existence of common guidelines and protocols across system facilities, extent and coordination of outcomes measurement activity, shared clinical support services, etc. Some measures of clinical “quality” (e.g., hospital-acquired infections) are not expressed in terms of their rates or levels per se, but rather in terms of the extent to which this information is collected in the same way and shared among system facilities.

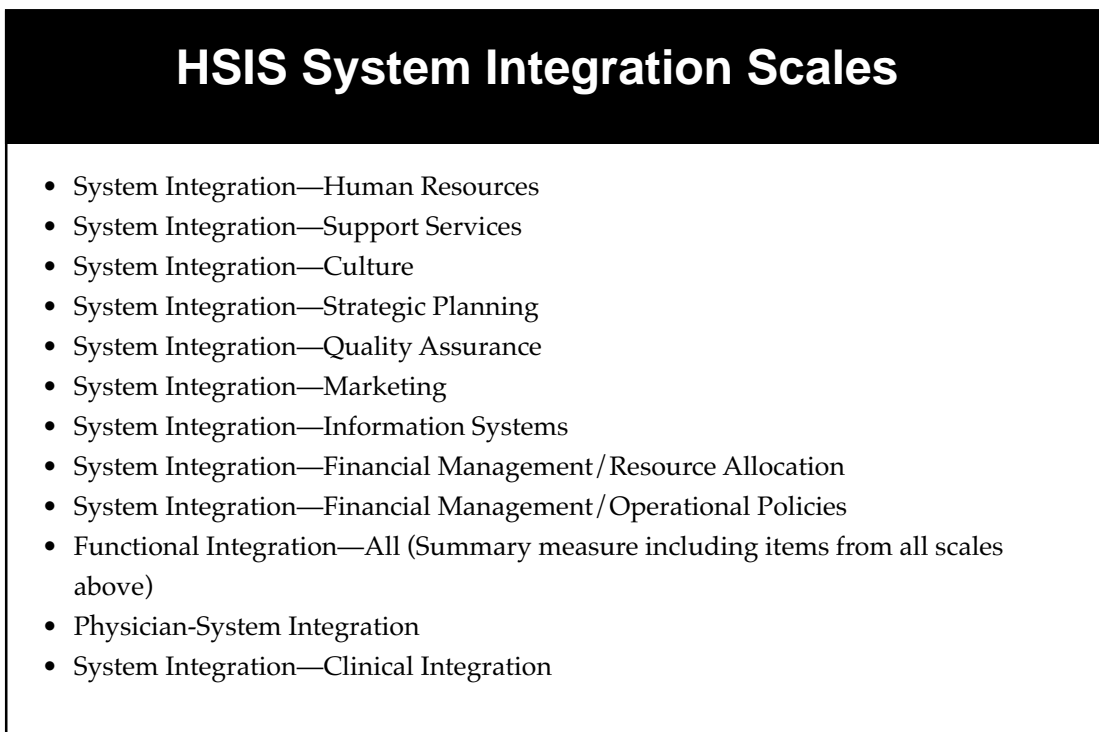
#### 4. Specific Measures

See Figure 5: HSIS System Integration Scales

#### 5. Current Status

The HSIS is essentially completed, and the investigator team continues to work with the data and publish articles that either compare levels of integration among the participating systems or identifies associations between various perceived or objective domains of integration. Researchers or organizations wishing to use the HSIS materials for not-for-profit purposes can do so by contacting Dr. Shortell's office and research staff at Northwestern University. The Northwestern investigators simply require a letter or e-mail describing the purpose for which the tools will be used, and ask that users pay a nominal fee to cover copying and mailing costs.

**Figure 5**



Source: Gillies, RR, Shortell, SM, Anderson, DA, et al. Conceptualizing and measuring integration: Findings from the Health Systems Integration Study. *Hospital and Health Services Administration*, 1993, 38, 467.

## D. Malcolm Baldrige Criteria

### 1. History and Purpose

The Malcolm Baldrige National Quality Award was established through an act of Congress in 1987, and has been administered through the Commerce Department. The award was designed to raise awareness about the importance of quality improvement, award organizations who had achieved special distinction in the area of quality improvement, and to encourage “benchmarking” or sharing of best practices among American companies.<sup>58</sup>

The award program consists of the development of a set of criteria for quality in major sectors of the U.S. economy (e.g., manufacturing, service industries, healthcare), then an annual competition in which companies go through a lengthy application process and review by a Board of Examiners to choose a very small number of actual award winners. The first awards were made in 1988; 22 organizations had won awards through 1994.

In 1995, a pilot program was established for healthcare organizations, recognizing both the unique criteria of quality in healthcare and the extent to which healthcare organizations had adopted CQI principles in the prior 7 to 8 years. The criteria were similar in basic organization to the more general criteria used for other kinds of organizations, but the criteria were tailored to reflect patient care, patient satisfaction, and healthcare improvements as unique “products.” The criteria are available for any organization to use for purposes of self-assessment; an annual review and awards process is conducted for organizations willing to prepare a formal application.

### 2. General Structure

The Baldrige Award criteria are similar to the SASI indicators in that they are open-ended rather than fixed and quantitative, they involve organizational self-assessment, and they are organized in major concept areas and specific subdivisions within those areas. There are specific point totals assigned to each major area and subarea, with a total of 1,000 points available if perfect scores are achieved in all. Judges who are experts in either healthcare or quality improvement review applications materials and conduct site visits in order to assign numeric scores to each area.

### 3. Main Sections or Important Subdivisions

There are seven major sections in the Baldrige Award criteria. They are:

1. *Leadership (90 points)*. Specific criteria that have to do with the effectiveness of leadership in the organization, clarity of vision, constancy of purpose, goal setting, and involvement of the organization as a community leader.
2. *Information and Analysis (75 points)*. Specific criteria that have to do with systems and capabilities for managing information, use of performance comparisons and benchmarking, and the use of information for strategic planning, resource allocation, and performance management.
3. *Strategic Planning (55 points)*. Criteria in this section assess the organization’s ability to analyze market needs and translate strategic plans into specific performance objectives, and alignment between clinical care or service enhancement plans and business plans.
4. *Human Resource Development and Management (140 points)*. Specific domains include human resource planning, staff development and training, credentialing, employee “ownership” of key processes and means of improvement, compensation, and employee satisfaction and well-being.

5. *Process Management (140 points)*. This section includes design of healthcare services, methods and capabilities for process redesign and improvement, systems in place for assessing quality of care and patient satisfaction, design of patient support services, and use of, or integration with, community-based health services.
6. *Organizational Performance Results (250 points)*. This section includes measures of patient healthcare outcomes, measures of process quality, community health benefits, administrative results, and results of accreditation or licensing reviews.
7. *Focus on, and Satisfaction of, Patients and Other Stakeholders (250 points)*. Measures include processes for market segmentation and analysis, methods for identifying patient needs and expectations, methods for resolving complaints and concerns, results of formal satisfaction surveys and other analyses, use of benchmarking and best practice analysis of satisfaction data, and other measures of patient or stakeholder loyalty to the organization.

The largest point total of any specific area is in the area of patient/stakeholder satisfaction. This emphasis is not surprising in light of the CQI emphasis on customer knowledge and designing processes to meet customer requirements. It does distinguish this measure set, however, from other measure sets in which patient satisfaction is just one of a large number of process and outcome measures with a stronger clinical emphasis.

#### **4. Specific Measures**

See Figure 6: 6.0 Organizational Performance Results (250 points)

Figure 6

## 6.0 Organizational Performance Results (250 points)

The Organizational Performance Results category examines performance and improvement in key patient healthcare areas, in patient care support service quality, and in community health services. Performance and improvement are assessed in administrative/business areas—productivity and operational effectiveness, supplier performance, and financial performance indicators linked to these areas. Also examined are performance levels relative to competitors and other institutions providing similar healthcare services.

### 6.1 Patient Health Care Results (80 points)

Summarize current trends and results of improvement efforts for key measures and/or indicators of patient health.

A     D     R  
 —     

#### Areas to Address

- a. For major healthcare services provided, give current levels and trends in key measures and/or indicators of patient health and the improvement of patient health.
- b. For the results reported in 6.1a, provide appropriate comparative data for other institutions providing similar healthcare services.

#### Notes:

(1) Results reported in Item 6.1 should reflect performance relative to specific key patient healthcare quality requirements. Such key quality requirements should relate closely to patient satisfaction and to standards of good practice. These requirements should be determined by factors in the organization's "Business Overview" and addressed in Items 7.1, 5.1, and 5.2.

(2) Data appropriate for inclusion are based upon:

- internal (organizational) measurements;
- data collected by the organization through follow-ups (7.2c) or surveys of patients;
- data collected by other organizations on the applicant's behalf.

(3) Measures and/or indicators (6.1a) of patient healthcare service quality depend upon setting and could include compliance with standard care patterns, unplanned revisits, nosocomial infections, immunization rates, low birth weight, length of stay, hospital readmissions, disease-specific and procedure-specific mortality, frequency of preventable acute episodes for chronic conditions, functional status, incidences of adverse drug interactions, hospital admissions per member/client/practice patient, coordination and continuity of care, and clinical outcomes. Appropriateness, effectiveness, and efficiency could be reported in Item 6.1 if considered as indicators of clinical outcome.

(4) Comparative data (6.1b) might include healthcare industry "best," data from similar institutions, and appropriate benchmarks. Such data might be derived from surveys, published and public studies, participation in indicator programs, or other sources.

Source: 1995 Health Care Pilot Criteria

## 5. Current Status

The Baldrige Award criteria are easily obtained from the Commerce Department. There are a number of private consultants or consulting firms available to help organizations through a self-assessment process as either a step prior to formal application or an end in itself. Far more organizations obtain the criteria and use them for self-assessment than actually go through the full process of applying for the award. Gathering and organizing information for either a self-assessment or a formal application is an extremely time-consuming process, but is judged worthwhile by many organizations as a means of organizing for improvement.

The 1997 award criteria (not specific to healthcare) have essentially the same seven main areas of evaluation and a similar set of criteria written in a more generic way to apply to many different kinds of commercial enterprises. Healthcare organizations could find either the 1995 healthcare criteria or the 1997 general criteria useful for self-assessment and possible benchmarking with peer organizations.

The balanced budget legislation for 1997 threatened to eliminate funding for expanding the Baldrige Awards to healthcare and education. At the time of writing of this report, the Senate version of the appropriation bill for the Commerce Department did not include expansion funding, but the House version did. The conference to iron out differences between the two bills had not yet occurred, and the House version had not yet been voted on and passed.

## E. CRISP

### 1. History and Purpose

The Consortium for Research on Indicators of System Performance (CRISP) was established in 1992 in response to publication in 1991 of Measures of Performance for Vertically Integrated Health Systems by the American Hospital Association. CRISP was established with 14 original members who shared an interest in the conceptual model and specific measures identified in the AHA report. The group's original goal was to test, refine, and implement as many of the measures included in the AHA report as possible, recognizing that many had not been used at that time for the purposes of comparing vertically integrated health systems.

The AHA report was the final report of the Ray Woodham Visiting Fellowship Program of 1990. The report described the unique organizational features and fundamental mission of vertically integrated systems, defined a number of domains in which performance measurement would identify the extent to which systems were meeting their mission goals, and provided a detailed list of 91 specific, quantitative measures of system performance.

The first year of consortium activity (1992) was devoted to "feasibility studies" of members' ability to collect data and report on the list of measures in the 1991 AHA report. Two-day site visits to member organizations were used to determine the extent to which medical records, scheduling, billing, quality assurance, and other existing databases could be used to obtain relevant information. Medical and administrative leaders were also interviewed to identify priority areas for measurement and obtain concurrence with overall project objectives.

CEOs of the member systems met in late 1992 to review results of the feasibility analyses and set priorities for data collection work in the following year. The members decided on 12 measures for a first round of data collection and analysis, with an additional six measures identified as "Tier II" measures that would receive either additional developmental work or some small scale pilot study effort to resolve issues of sampling, coding, or data analysis/interpretation.

In 1993, the first round of data collection took place, and comparative performance reports were issued to member systems by the end of the year. Some improvements to sampling and data collection methods were made for the following year, and there have been annual cycles

of data collection and reporting since then. Some of the original 12 measures have been dropped because of either insufficient interest or technical difficulties that could not be resolved; other measures have been added or modified, including many of those that were on the original “Tier II” list.

The CRISP consortium is still focused on testing and refining clinical and administrative performance measures, and has submitted materials to JCAHO for review and inclusion in the JCAHO’s lists of approved measure sets for both hospitals and healthcare networks in its ORYX initiative (see JCAHO section below). A core project staff based at Henry Ford Health System in Detroit provides daily coordination for the project; an annual meeting in Detroit provides the project steering committee with a chance to review activity and set priorities for the upcoming year. Work on CRISP has been supported by annual contributions from the member systems.

## 2. General Structure

The current CRISP measure set is organized around three specific clinical conditions (soon to be four), and a set of administrative and community benefit indicators that are not disease-specific. Within each clinical condition or administrative area, there are several specific performance measures that are derived from widely accepted clinical guidelines, published research findings, or prior performance measurement activities.

In the clinical measures, there are detailed instructions for patient inclusion/exclusion and for obtaining relevant data from either administrative databases or patient medical records. In two of the areas, surveys are included to assess patient satisfaction and functional health status outcomes. Detailed instructions for survey administration are provided. A fixed sample size of relevant cases is identified (e.g., 200 most recent live births at system hospitals) and specific time periods during the year are identified during which cases can be sampled.

Data are obtained by employees of member systems themselves, using whatever data sources are available. Data collection manuals define the appropriate procedure or diagnosis codes and definitions of key variables; the systems themselves have considerable flexibility in identifying data sources and assembling databases for transmission to the project offices in Detroit.

For the administrative measures, a similar set of detailed data collection instructions and definitions are provided, but the measures do not require the identification of individual patients or patient records. Data collection forms are filled out by administrative staff using existing information from management and financial reports.

## 3. Main Sections or Important Subdivisions

The 1997 CRISP measure set includes the following major areas:

- *Breast Cancer Care*. Measures include stage of cancer at diagnosis, “sleepless nights” (time from abnormal mammogram to definitive diagnosis and from diagnosis to treatment), and proportion of positive biopsies.
- *Acute Myocardial Infarction Care*. Measures include several “timeliness” measures reflecting the immediate response of the system to a new or suspected AMI (e.g., time from arrival at facility to administration of thrombolytic if given), provision of recommended therapy (e.g., percent of patients with appropriate clinical indications given thrombolytic), use of recommended medications during admission and at discharge, and measures of patient satisfaction and health status three months post-hospital discharge.

- *Prenatal Care–Birth Outcomes.* One subset of measures includes low birth weight and early gestational age as outcome measures and includes a detailed set of process of care measures that reflect the early initiation and provision of prenatal care services according to generally accepted guidelines. Cesarean section and VBAC rates are included as part of this subset. The other subset focuses on the management of imminent preterm birth, particularly in terms of administration of corticosteroids to prevent respiratory complications in neonates.
- *Depression (Under Development).* A technical work group met in the late summer and fall of 1997 to choose a set of performance measures reflecting the systems' screening for, and management of, depression in primary care settings. Pilot data collection should begin in 1998.
- *Health Behaviors.* A standardized health behavior survey (Prevention Index) is administered to samples of either health plan enrollees or patients with ongoing primary care relationships with system-affiliated physicians.
- *Community Benefit.* Three specific measures—amount of charity care, underpayments in public programs (Medicare and Medicaid), and dollar value of community health programs—are analyzed both separately and in combination, using overall system expenses as a denominator.
- *Administrative Expense.* System expenses in a list of specific administrative expense categories (e.g., human resources, planning and marketing, billing, etc.) are summed and expressed as a percentage of total system expenses.
- *Financial Performance.* Measures of profitability and debt service coverage ratio are included in this section.

#### **4. Specific Measures**

See Figure 7: Current CRISP Indicators and Processes and Outcomes Being Measured

Figure 7

## Current CRISP Indicators and Processes and Outcomes Being Measured

### Breast Cancer: Project I

- Stage at diagnosis. NOTE: Breast cancer is a relatively low-incidence condition. From CRISP's experience and others, unless systems have very large populations (300,000+) systems will probably not find this indicator useful for either internal or external purposes

### Breast Cancer: Project II

- Proportion of positive breast biopsies
- "Sleepless nights": Time between significant events such as abnormal mammogram, biopsy, definitive diagnosis, and treatment
- Rates of breast conserving therapy and other procedures/treatment
- Pilot testing functional status and satisfaction (3-6 months) post-treatment initiation
- Estrogen receptor test rates

### Acute Myocardial Infection

- Time to thrombolytic or PTCA
- Medication utilization at discharge and two months postdischarge (aspirin, beta blockers, ACE inhibitors, calcium channel blockers)
- Timing of aspirin for patients with AMI
- Inpatient mortality and 60-day postdischarge mortality following AMI
- Patient satisfaction and functional status at 2 months postdischarge
- Appropriate treatment for elevated cholesterol levels

### Neonatal Health: Project I

- Utilization of prenatal care initiation and visits
- Cesarean section rates
- Rate of vaginal birth after previous cesarean section (VBAC)
- Low birth weight rates
- Very low birth weight rates
- Preterm birth rates (less than 37 weeks gestation)

### Neonatal Health: Project II

- Use of antenatal corticosteroids in cases of preterm birth (24-36 weeks gestation)
- Rate of major infant morbidity at 28 days postbirth
- Infant mortality rate at 28 days postbirth
- Low Apgar score in preterm infants (24-36 weeks)
- Length of stay of mother and infant
- Length of NICU stay for infant

### Community Benefit

- Proportion of system expense devoted to charity care
- Underpayment in public programs
- Value of community benefit programs

### Administrative Expense

- Percentage of total system expenses that are attributed to administrative expenses

### Financial Performance

- Debt of service coverage ratio
- Profitability, subcategorized by (a) total margin, and (b) operating margin

## 5. Current status

CRISP currently has nine active members. There were data collection cycles on virtually all areas of measurement in 1997; results will be analyzed and reported in the last part of the year. The measure set remained relatively stable into 1998, with the planned addition of a measure set for management of depression in primary care. The group will focus its efforts on those measure sets in which further development work can be supported by external grants or contracts, or through special, targeted contributions by member organizations. A single, global membership fee covering work on the entire set of indicators will not be continued into 1998. It is possible that the specifications and data collection instructions for some measures may be licensed to other JCAHO-approved measurement systems.

Descriptions of CRISP measures, and the comparative performance reports, are essentially in the public domain. They are available from the CRISP project office at Henry Ford Health System on request. Member health systems are identified only by code number on the reports, so that it is not possible to identify specific organizations without access to the code system. Members receive the information necessary to identify their own system in a table or figure showing comparative performance, but they are not able to identify other individual systems.

Although the CRISP measure set recently gained JCAHO approval as part of the ORYX initiative, it had not been decided at the time of writing of this monograph whether members would support CRISP functioning as a performance measurement "vendor." The group's history is as a voluntary research consortium rather than a "vendor" with "clients" who use the data management services of the vendor to meet accreditation or purchaser requirements. On the other hand, pressures to focus on required measurement will make it difficult for members to commit to voluntary efforts like CRISP if doing that work does not meet JCAHO accreditation requirements. The consortium is focusing its current efforts on seeking external grant funding for specific research questions related to performance measures in the specific clinical areas (e.g., How do variations in process of care measures at the system level relate to variations in outcome measures at the same level?).

## F. Foundation for Accountability (FACCT)

### 1. History and Purpose

FACCT was established in 1995 under the leadership of Dr. Paul Ellwood and the Jackson Hole Group and many of the same large private employers (e.g., GTE, Xerox, Digital) who were involved in the development and revision of HEDIS. The group was expanded to include representatives of large public employers and purchasers, including the HCFA and two large state employee health benefits programs. Healthcare plans and providers were not active participants in the early stages of FACCT, but have been involved more recently in discussion of implementation of FACCT measure sets.

The main purpose of FACCT has been to advocate for the collection and reporting of patient outcome and satisfaction data, and to define a standard set of measures for that purpose. FACCT was formed, in part, to maintain an interest in the parts of the Clinton 1993 Health Care Reform proposals having to do with competing health systems and competition on the basis of health outcomes, and, in part, out of concern that HEDIS had not gone far enough in including outcome measures rather than process of care/service provision measures. Both motivations reflected a basic philosophic premise that health plans and provider groups could, and should, be held accountable for health status outcomes.

FACCT has released several sets of recommended performance measures for specific clinical conditions. These measure sets were the results of commissioned work by academic experts in the field and consultation with representatives from plans and provider organizations who have had experience working in those clinical areas. There have been discussions of local market area pilot studies, and funding of at least one pilot study of FACCT measures by the Agency for Health Care Policy and Research through its Q-SPAN quality research initiative.

## **2. General Structure**

The FACCT measure sets represent a “consensus” among selected experts in the field on the best available measures of both process of care and outcomes for patients with a specific illness. The measure sets typically include lists of demographic and clinical variables that should be included in the data collection for purposes of case-mix or severity adjustment. Detailed sets of codes, definitions, and data collection recommendations are provided for each measure set. Many of the measure sets (e.g., depression) make extensive use of patient surveys as a primary data source.

## **3. Main Sections or Important Subdivisions**

The FACCT measure sets released in 1996 are organized around specific diseases or medical conditions, including:

- diabetes
- depression
- breast cancer
- health maintenance in an enrolled population

An asthma measure set was developed but not released in 1996 due to concerns about the technical soundness of the measures and the ability to use them to make valid comparisons among plans or providers.

## **4. Specific Measures**

See Figure 8: Foundation for Accountability (FACCT) Performance Measures

**Figure 8**

<b>Foundation for Accountability (FACCT) Performance Measures FACCT Measures to Assess Satisfaction With Care and Services in a Health Plan</b>		
Quality Measure	What Each Measure Tells You	How Each Is Measured
<b>Satisfaction</b>		
Getting Needed Services	How satisfied members are in getting an appointment for care and getting the care they need.	Yearly patient/member questionnaire
Skill of Care Providers	How satisfied members are with the skill of their care providers.	Yearly patient/member questionnaire
Having Choices of Doctors and Other Care Providers	How satisfied members are with the choices they have to select and change their doctors and other care providers.	Yearly patient/member questionnaire
Good Results from Your Treatment	How satisfied members are with the results of their care.	Yearly patient/member questionnaire
Recommend Plan to Others	How many members are satisfied enough with their health plan to plan to stay in the health plan and recommend it to others.	Yearly patient/member questionnaire
Overall Quality of the Health Plan	How satisfied members are with the overall quality of their health plan.	Yearly patient/member questionnaire
<b>FACCT Measures to Assess Quality Care For People’s Health Risks in a Healthcare Organization</b>		
Quality Measure	What Each Measure Tells You	How Each Is measured
<b>Steps to Good Care</b>		
Helping Smokers Quit	Do the doctors and other care providers offer help to smokers to stop smoking.	Smokers complete a yearly questionnaire.
Care Providers’ Awareness of Members’ Health Habits	Does the healthcare organization survey its members to learn the extent of their poor health habits.	Check that the healthcare organization surveys its members/patients about their health habits.
<b>Results</b>		
Quit Smoking	How many smokers quit smoking during the past year.	Smokers complete a yearly questionnaire.

–Continued on next page

Figure 8 (continued)

<b>FACCT Measures to Assess Quality of Breast Cancer Care in a Healthcare Organization</b>		
<b>Quality Measure</b>	<b>What Each Measure Tells You</b>	<b>How Each is Measured</b>
<b>Satisfaction</b>		
Testing for Cancer Regularly	How many older women have a yearly mammogram to test for breast cancer.	Doctor's billing or claim records for yearly mammography exams provided to women ages 50-75.
Finding Cancer Early	How many patients' breast cancers were detected early when the chances for recovery are greater.	Centralized cancer records contain patient histories including extent of cancer when it was first found.
Making Informed Decisions About Treatments	Do patients with less advanced breast cancer receive necessary information about their surgery options.	Patients, 3 months after diagnosis, complete a questionnaire.
Receiving the Most Effective Treatments	How many patients, with less advanced cancer, undergo conservative breast surgery instead of full breast removal and did they receive the needed radiation treatment after surgery.	Centralized cancer records contain patient histories including types of treatment for women with less advanced cancer.
<b>Satisfaction</b>		
Good Communication	How satisfied patients are with their communication with doctors and nurses, their involvement in treatment decisions, and the timeliness of getting test results.	Patients, 3 months after diagnosis, complete a questionnaire.
Getting Needed Services	How satisfied patients are with being able to see specialists and getting support services.	Patients, 3 months after diagnosis, complete a questionnaire.
<b>Results</b>		
Five-Year Disease-Free Survival	How many patients are treated successfully without a return of the cancer	Centralized cancer records contain patient histories including extent of the cancer 5 years after it was found.
Coping with Disease	How well patients continue their routine activities and cope with the cancer and its treatment	Patients, 1 year after diagnosis, complete a questionnaire.

**Figure 8 (continued)**

<b>FACCT Measures to Assess the Quality of Major Depressive Disorders Care in a Healthcare Organization</b>		
<b>Quality Measure</b>	<b>What Each Measure Tells You</b>	<b>How Each is Measured</b>
<b>Steps to Good Care</b>		
Strong Links with Care Providers to Get Needed Care	How many patients feel sufficiently connected to their care providers and are able to get the care they need over time.	Patients, 6 months after diagnosis, complete a questionnaire.
<b>Satisfaction</b>		
Being Treated with Respect	How satisfied patients are with the respect and attention given them by their care providers.	Patients, 6 months after diagnosis, complete a questionnaire.
Skill of Care Providers	How satisfied patients are with the skill and responsiveness of their care providers and the results of their treatment.	Patients, 6 months after diagnosis, complete a questionnaire.
<b>Results</b>		
Recovery Within 6 Months	How many patients' conditions improve emotionally after treatment.	Patients complete a questionnaire when they are diagnosed and again, 6 months later.
<b>FACCT Measures to Assess Quality Care of Diabetes Care in a Healthcare Organization</b>		
<b>Quality Measure</b>	<b>What Each Measure Tells You</b>	<b>How Each Is Measured</b>
<b>Steps to Good Care</b>		
Checking for Foot Problems Regularly	Whether patients have their feet checked to avoid losing a foot to disease.	Patients complete a questionnaire yearly.
Checking Blood Sugar Regularly	Whether patients have their blood sugar checked to avoid more serious illness	Laboratory billing or claims records.
Getting Yearly Eye Exams	Whether patients have their eyes checked at least yearly to detect problems that lead to blindness.	Claims records.
<b>Satisfaction</b>		
Getting Needed Services	How satisfied patients are with ability to see specialists and getting support services.	Patients complete a questionnaire yearly.
<b>Results</b>		
Coping with Disease	How well patients continue their routine activities and cope with diabetes and its treatment.	Patients complete a questionnaire yearly.
Keeping Blood Sugar in Control	How many patients keep their blood sugar at levels that help avoid more serious illness.	Laboratory test results obtained from the lab or patient's medical record.
Keeping Cholesterol at Safe Levels	How many patients keep their cholesterol at levels that help avoid heart disease.	Laboratory test results obtained from the lab or patient's medical record.
Avoiding Hospital Stays	How many patients can manage their diabetes at home, with their doctor's help, and avoid being hospitalized.	Hospital billing or claim records.

Source: *FACCT Releases Its First Quality Measures for Health Care Performance* (press release). Portland, OR: Foundation for Accountability, June 1996. Reprinted with permission.

## 5. Current Status

The FACCT measures have received a great deal of attention and visibility because of the prominence of Dr. Ellwood and the sponsoring organizations and because of the excellent descriptive materials developed by FACCT staff. At the time of this writing (late summer/fall 1997), there are no published reports of actual FACCT data collection and analysis. Many of the specific measures have already been used in the context of research projects or other types of comparative performance reports, so there is some confidence in the probable value and “face validity” of the measures. However, the extensive revision of HEDIS and CRISP measures, carried out as a result of field testing, has not yet occurred for the FACCT measures.

## G. JCAHO Criteria

### 1. History and Purpose

In the mid-1980s, the Joint Commission for Accreditation of Healthcare Organizations (JCAHO) announced its “Agenda for Change” initiative. This initiative recognized the need for JCAHO and the organizations it accredits (primarily hospitals) to move beyond the existing accreditation standards based largely on structural and procedural standards to another approach based more on actual measurement and comparison of organizational performance.

One key component of the Agenda for Change was a clinical measure set known as the “IMSystem,” described for the first time in 1986. This measure set was designed to capture key aspects of quality of care for hospitals, expressed in quantitative terms suitable for comparisons and relative rankings. The measures were designed to be very clinically detailed and specific, and to capture performance of most major clinical areas and departments in a typical hospital (e.g., anesthesiology, oncology, cardiology, obstetrics, etc.).

The IMSystem measures were “alpha tested” in a volunteer set of 34 hospitals between 1987 and 1993 (different measure sets were tested by different hospitals in different years). A “beta test” phase involving a total of 451 hospitals from 1990 through 1995 resulted in some refinement and more major revision of the measure sets.

Test hospitals began expressing concerns to JCAHO about both the burden of data collection and the value of the information for internal use (and possible use in accreditation reviews). These concerns escalated to the point of a significant crisis for JCAHO in 1994 and 1995. JCAHO had made a major public commitment to the IMSystem as the backbone of its future accreditation process, and the objections were serious enough to force the major constituent organizations of JCAHO (AHA, AMA, ADA, American College of Surgeons, American College of Physicians) to consider either withdrawing their support of JCAHO itself or forcing a major change of direction.

In the meantime, the more traditional JCAHO accreditation criteria remained in place and continues to. These criteria emphasize adherence to a detailed set of standards and norms for organizational, infrastructure, and procedural dimensions of hospital operations (e.g., presence and active operation of key quality assurance committees, cleanliness of facilities, adequacy of library facilities, etc.).

In response to the concerns expressed by both hospitals and the organizational sponsors, JCAHO backed off from its stated plans of making the IMSystem mandatory for 1996, and adopted an alternative strategy that involved a “call for indicators.” JCAHO’s new approach was to receive, review, approve, and distribute information about other performance measurement systems, and offer hospitals (and health networks) the option of using one or more measure sets on the “approved” list.<sup>59</sup>

This approach has proven far more successful, both in terms of hospital acceptance and in terms of response to the two rounds of “calls for indicators” that have been issued. Sixty measure sets are on the approved list for hospitals and five measure sets are on the approved list for health plans and networks. (An expanded list of approved measure sets was released in November, 1997, based on applications from candidate measurement systems to JCAHO in the summer of 1997).

JCAHO’s purpose is still to move the field in the direction of quantitative, comparative measures of process and outcome rather than complete dependence on structural measures of quality, but the new approach offers much more flexibility and convenience for hospitals than did the original approach of the mandated IMSsystem measures. The JCAHO’s new program is called ORYX; a subset of the original IMSsystem measures remain an approved option under the name “ORYX Plus.”<sup>60</sup>

For networks, the JCAHO indicator set functions as a complement to a set of standards and criteria used for network accreditation. The network accreditation program began in 1995, and has been used to accredit 19 networks as of mid-1997.<sup>61</sup> The accredited networks are a mix of HMOs, specialized health services networks (e.g., Health Services for Children with Special Needs), and more comprehensive delivery networks. For the HMOs, JCAHO network accreditation can be seen as an alternative to, complement to, or perhaps a competitor to, accreditation by the NCQA. For IPAs, PHOs, integrated delivery systems, and more specialized care networks, the JCAHO network accreditation program offers the opportunity for recognition and accreditation of an organizational level that would otherwise slip between the cracks of existing licensing/accreditation programs.

## **2. General Structure**

The ORYX performance measures program now consists of 60 approved measure sets for hospitals and five approved measure sets for health plans and delivery networks. Each measure set has its own sponsoring organization or vendor and its own detailed instructions and procedures for data collection, analysis, and reporting. Most of the measure sets already have the active participation of multiple hospitals or hospital systems, allowing individual hospital performance to be compared to norms and benchmarks from peer organizations.

Hospitals can choose one or more of the approved measure sets for use. The 1997 requirement is simply to choose a measure set; in 1998, hospitals must be actively involved in a set of measures that, taken together, cover care provided to at least 20% of the hospital’s patients. The sponsoring organizations or vendors for the measure sets are obliged to send quarterly performance reports, starting in March of 1999, to HCFA that identify individual hospitals and their standing in comparative performance analyses.

The JCAHO network accreditation standards are similar to the standards and criteria for hospital accreditation in that they specify certain structural and procedural characteristics of organizations that JCAHO views as essential to the provision of high-quality care. For networks, specific criteria have been written in areas such as: systemwide integration, coordination of care, quality improvement, credentialing and privileging processes, etc. JCAHO has recently begun a “Charter Network Accreditation Project” to tailor the accreditation criteria and review process to integrated delivery networks (IDNs).<sup>61</sup> There is thus some separation of criteria and review process for health plans and integrated delivery networks; a separate accreditation manual exists for preferred provider organizations (PPOs).

## **3. Main Sections or Important Subdivisions**

The key distinction in the ORYX performance measurement system is between measures of hospital performance and measures of plan or network performance. The hospital measures are required for JCAHO hospital accreditation, so there is considerable pressure for hospitals to choose a measure set (if they have not already done so) and become active in data collection. Hospitals currently involved in measurement activities not on the approved list are forced either to switch or to request that their measurement system vendor seek JCAHO

approval by 1998. The health plan and network measures are required for JCAHO network accreditation on a slightly more relaxed time schedule, but very few organizations (15 or less) have actually sought JCAHO accreditation at that level. NCQA is still by far the dominant entity for health plan accreditation, and the HEDIS measure set is the one required by NCQA. There is no consensus about network accreditation, either in terms of the necessity for it or the proper organization to do it.

The hospital measure set list includes 60 measure sets that span a wide range of clinical specificity and detail. The APACHE III system is an example of a very clinically detailed system focused on measures of ICU care.<sup>62</sup> On the other hand, the Maryland Hospital Association measure set is a broader, more comprehensive set of measures covering nearly all departments and patient groups.<sup>63</sup>

The health plan and network measure set includes only five sets: HEDIS, FACCT, and three measure sets with a more specific patient population in mind, either nursing home care, home healthcare, or rehabilitation.

See Figure 9: JCAHO Approved Network Measure Sets

The JCAHO criteria and standards for network accreditation include a large number of specific criteria under the major headings summarized in Figure 10. The major areas and key concepts being addressed are:

- *Rights, Responsibilities, and Ethics.* Criteria have to do with policies and procedures about confidentiality of information, patient/family rights to appeal of decisions, do-not-resuscitate orders, and access to information.
- *Continuum of Care.* Criteria address scope of services, access, fit between services and member needs, information flow across sites and providers, and appropriate mix of providers.
- *Education and Communication.* Criteria address the existence and quality of information provided to members/patients about covered benefits, use of services, appeal procedures, self-care, disease prevention, etc.
- *Leadership.* Criteria describe the management structure of the network, its ownership and financial structure, and the mechanisms in place for setting fiscal and clinical policies. Criteria about policies for working with contractors are also included.
- *Management of Human Resources.* Criteria focus on policies and procedures for credentialing and privileging of practitioners, verification of training, adequacy of numbers of practitioners to meet member/patient needs, and periodic review of the above.
- *Management of Information.* Criteria address design of information systems, security of information, use of standard coding conventions, data quality control, and use of data for comparative quality measurement.
- *Improving Network Performance.* Criteria determine the extent to which the network is committed to quality improvement in both clinical and administrative domains, involves clinicians in improvement efforts, and has measurement capabilities in place to assess improvement.
- *Health Promotion and Disease Prevention.* Criteria address the network's ability to define and provide primary, secondary, and tertiary preventive services, appropriate to the needs of its members/patients.

**Figure 9**

<b>JCAHO Approved Network Measure Sets</b>		
<b>Organization</b>	<b>Focus</b>	<b>Contact</b>
Joint Commission on Accreditation of Healthcare Organizations (JCAHO)	Acute Care Hospitals (IMSystem)	Department of Indicator Measurement IMSystem Help Desk One Renaissance Blvd. Oakbrook Terrace, IL 60181
The Foundation for Accountability (FACCT)	Networks/ Health Plans	220 NW 2nd Avenue Suite 725 Portland, OR 97209 (503) 223-2228
National Committee for Quality Assurance (NCQA)	Networks/ Health Plans (HEDIS 3.0)	Dept. of Planning and Development Suite 500 2000 L Street NW Washington, DC 20036 (202) 955-3520
University of Colorado Health Sciences Center	Home Care	Center for Health Services Research 1355 South Colorado Blvd. Suite 306 Denver, CO 80222 (303) 756-8350
University of Wisconsin, Madison	Long-Term Care	Center for Health Systems Research and Analysis 610 Walnut Street Room 163 WARF Building Madison, WI 53705 (608) 263-4523

**Figure 10**

<b>Summary of Network Accreditation Criteria</b>
<ul style="list-style-type: none"> <li>• Rights, Responsibilities, and Ethics</li> <li>• Continuum of Care</li> <li>• Education and Communication</li> <li>• Leadership</li> <li>• Management of Human Resources</li> <li>• Management of Information</li> <li>• Improving Network Performance</li> <li>• Health Promotion &amp; Disease Prevention</li> </ul>

#### **4. Specific Measures**

Given the large number of measure sets in the JCAHO ORYX system, space does not permit either a listing or detailed discussion of each specific measure. There are thousands of them. Many are similar or even identical from measure set to measure set. The JCAHO web site (<http://www.jcaho.org>) provides a relatively detailed description of each approved measure set and at least a partial listing of the specific performance measures included.

#### **5. Current status**

JCAHO has recovered extremely well from its 1994-1995 "crisis" and is playing a very active and positive role in the continued development of performance measurement systems. JCAHO approval has become extremely important for all measurement systems that wish to maintain a client base of hospitals.

This flurry of activity is apparent for hospital performance measures; much less activity is present on the health plan and network part of ORYX. The CRISP measure set has been submitted for review and possible approval; HEDIS and FACCT are already approved. Unless JCAHO becomes seen as a legitimate competitor to NCQA in the area of health plan accreditation, and unless network accreditation becomes mandatory or important for contracting purposes for PHOs, PSO, or IPAs, the attention and activity will continue to be primarily in the area of hospital performance measures and accreditation.



## IV. Comparison / Evaluation of Measurement Systems

Some of the strengths and weaknesses of the performance measurement systems have been identified in the preceding descriptions. In this section, the systems will be evaluated according to the major criteria for measures that were discussed earlier.

### A. Is the System Designed Around a Clear Purpose and Target Audience?

The SASI tool, HSIIS measure sets, and Malcolm Baldrige criteria are all designed primarily for internal self-assessment by health systems or health networks. The SASI tool is noteworthy for its specific focus on this purpose. The HEDIS, CRISP, JCAHO, and FACCT measure sets have all been designed for a combination of internal improvement, external reporting/consumer choice, and accreditation purposes. Within this set, HEDIS is probably most clearly defined in its use for reporting health plan performance to employers.

The target audience for reports from SASI, HSIIS, and Malcolm Baldrige would probably be network leaders or other key stakeholders who have an active role in managing and planning network activities. The primary target audiences for HEDIS and JCAHO reports are purchasers and accrediting bodies who are evaluating plans, networks, or individual hospitals. Individual consumers are also an important audience. Plan, network, and hospital leadership can use reports for quality improvement and benchmarking purposes. Health system leaders are the primary audience for CRISP reports, since participating organizations are not identified except by code number. FACCT data could be used by several different target audiences, including purchasers, individual consumers, and plan or provider network managers, depending on the details of FACCT implementation.

The three major purposes of comparative performance reports are:

1. *Internal Improvement or Benchmarking.* All of the measure sets could conceivably be used for this purpose. Some of the specific measure sets on the JCAHO list and some of the CRISP or FACCT measures are perhaps best suited for this purpose because they are clinically specific, oriented to single clinical processes (e.g., administration of thrombolytic), and can be collected on a case-by-case basis in nearly “real time.” Within these measure sets, measures of process rather than more global measures of longer-term outcome (e.g., 5-year survival rate for breast cancer) would be preferred. Many of the HEDIS measures are linked closely to processes that can be improved, but their structure as rates means that many individual cases over a period of time must be aggregated to produce a single data point on a time line. Most process improvement efforts have a more rapid “cycle time” than rate-based measures can support.

The SASI, HSIIS, and Malcolm Baldrige criteria may include measures suitable for internal improvement use, but the more open-ended format or use of opinion surveys makes comparisons with other organizations less direct, and the link between indicators and specific clinical processes less clear. A survey measure of “clinical integration” HSIIS, for example, does not lead directly to either a specific clinical process or a potential process improvement.

The “benchmarking” concept includes a range of activities that serve the purpose of identifying best practices and comparing one’s own organizational performance to that standard. For that general purpose, all of the measure sets can potentially be used.

2. *Consumer Choice.* The HEDIS, JCAHO, and FACCT measure sets stand apart from the others in this regard, since they were all designed specifically for this purpose, at least at the level of “corporate” consumers. None of the measure sets were designed specifically to meet the needs of individual consumers making choices among plans and providers, and they all include some information that consumers will find irrelevant or confusing.<sup>64,65</sup> Nevertheless, individual consumers choosing among health plans will find some value in HEDIS data; consumers choosing among hospitals will find some value in many of the measure sets approved by JCAHO. The FACCT measures are potentially useful and unique because of their inclusion of patient outcomes data; their actual utility remains to be demonstrated.

The CRISP, SASI, HSIS, and Malcolm Baldrige measures or criteria are not nearly as useful, either because they do not actually identify organizations by name (CRISP) or because they are not designed to produce comparative reports (SASI, HSIS, Malcolm Baldrige). The CRISP measures could be useful for consumer choice if used in a different context where network identities were released.

3. *Accreditation.* Participation in HEDIS is part of the NCQA accreditation review process for health plans; participation in one of their approved measure sets is now required for JCAHO accreditation of hospitals and will be required in the near future for accreditation at the plan or network level. Either the HEDIS or FACCT measure sets are currently acceptable to JCAHO for network accreditation.

## **B. Are Measures Defined in Quantitative, Objective Terms?**

The HEDIS, CRISP, and FACCT measures are generally both quantitative and objective. Some of the specific measures or indicators that are based on patients’ or members’ survey responses (e.g., satisfaction) may not be purely objective, but the data can be used in that way. Many of the HSIS system performance measures are both quantitative and objective (e.g., number of physicians admitting 10 or more patients to system hospitals); the system integration measures are quantitative but subjective, since they are based on opinions of various system stakeholder groups.

The SASI tool and the Malcolm Baldrige criteria allow for the inclusion of quantitative, objective data, but the format does not strictly require that kind of information, and the basic structure of the indicators is more open-ended, text-oriented, and subjective. The “Global Measures” section of the SASI tool does include items that are more quantitative in nature.

## **C. Do Measures, Taken Together, Provide a Comprehensive “Picture” of the Network?**

The SASI tool, HSIS measures, and Malcolm Baldrige criteria are superior to the others in this regard for two reasons. First, at least within their stated domain (e.g., system integration for HSIS), the measure sets are specifically designed to describe the network/system in its entirety. Second, the network is the organizational entity that the designers of the measures had in mind.

The other measure sets have clear limitations on this dimension. HEDIS is a fairly comprehensive measure set, but there are clearly many clinical areas left out, and even HEDIS 3.0 includes very few outcome measures. HEDIS is designed to assess health plans, and is not necessarily suitable for assessing delivery networks within plans. Most of the measures in HEDIS 3.0 were designed for commercially insured (i.e., under 65) member populations, so the specific measures are not well suited for chronic care networks. The

CRISP measure set is even more “sparse” since member organizations have chosen to focus their data collection and analysis activities on only a selected set of clinical and administrative measures. CRISP does focus on delivery networks rather than plans as the unit of analysis, but has not identified chronic care networks as a specific area of focus. The clinical conditions currently active in CRISP are not well aligned with the priorities of chronic care networks. It is difficult to make a general statement about the JCAHO measure sets, since there are so many of them and their contents are so varied. Their focus on hospital care, though, makes them inherently limited in application to a network with ambulatory, home health, rehabilitation, and other non-inpatient components.

## **D. Does the System have Documented Reliability, Validity, and Sensitivity?**

The HEDIS, CRISP, and JCAHO measure sets have gone through extensive testing and have passed at least some basic tests of their psychometric properties. The measure sets include many specific measures (e.g., SF-36 survey) that have been through even more extensive and rigorous testing by their original developers. The sensitivity of the measures to either process improvements or more general organizational change over time is less clearly understood, but is becoming more clear as the measures are used for comparative purposes.<sup>66,67</sup>

The FACCT measure sets include specific measures with good and well-documented psychometric properties; their performance in comparing health plans or providers and informing consumer choice has yet to be demonstrated.

The SASI and Malcolm Baldrige criteria allow the inclusion of measures with proven reliability, validity, and sensitivity, but the format of these instruments or indicator sets does not require this. The HSIS measures were pretested in one organization before being used in the HSIS and appear to be psychometrically sound. The sample size in the HSIS, though, is far too small to have allowed for detailed analysis of the properties of the measures at the network or system level.

## **E. Are Adjustments Made for Case Mix or Disease Severity?**

The CRISP clinical measures, the proposed FACCT measures, and many of the specific measures or measure sets on the JCAHO list have adjustments for case mix or severity included. HEDIS measures generally do not, although development of case mix adjustment methods has been identified by NCQA as an important area for future HEDIS development.<sup>68</sup>

The concept of case mix or severity adjustment does not apply nearly as well to the SASI, HSIS, and Malcolm Baldrige criteria, where the “unit of measurement” is the network itself, not individual patients or patient groups within the network.

## **F. What is the Appeal to Users?**

### **1. Face Validity: Basis in Peer Experience for Lay Users**

All of the measure sets reviewed here have a certain level of “face validity” that has been part of their current level of acceptance and visibility. Most are undergoing some amount of continual revision and improvement so that the most useful measures remain, new measures are added that have potential value, and less useful measures are either changed or eliminated.

One criterion for judging appeal to users is the time and effort required to obtain the raw data and package it in a form suitable for use. HEDIS 2.0 and 2.5 were designed to draw on

existing claims databases with an option of using medical record abstraction. The 1994 pilot study demonstrated significant disagreement between the two approaches, so many plans have asked providers to gather information from medical records sources to generate more complete and accurate HEDIS reports.<sup>69</sup> HEDIS 3.0 requires a significantly greater investment of time and effort, and some of the initial reactions to the release of the specifications in 1996 were complaints about the amount of resources required for data collection.<sup>70</sup>

Many of the CRISP measures require significant manual abstraction of medical records, so the same concerns apply about the amount of staff time required for data abstraction. The consortium has estimated that one full-time employee is required for data gathering if the organization commits to the full list of CRISP measures.<sup>71</sup> The FACCT measures are similarly data- and labor-intensive, but the actual amount of time required will not be known until the measures are actually used.

SASI, HSIS, and Malcolm Baldrige assessments can all be expected to be extremely time-consuming, since part of the process involves teams of individuals in the network to be involved in interviews, group meetings, survey completion, or other forms of data collection. Assembling the information, preparing materials for outside reviewers, and participating in site visits all add to the time commitment required for a Malcolm Baldrige application.<sup>72</sup> NCCC staff report that the SASI tool can be used in as short a time as one day (with some preparation work) or as long a period as six to nine months as part of a committee process.

## **2. Clinical Detail for Acceptance by Clinicians**

HEDIS, CRISP, FACCT, and most of the JCAHO measure sets include measures that are defined in clinically precise terms, usually with the involvement of clinicians as members of an expert panel. SASI indicators, HSIS measures, and Malcolm Baldrige criteria are generally not clinically detailed, since they reflect organizational structure and dynamics rather than care processes for specific diseases. The three measure sets, though, do have places for inclusion of clinically specific measures as illustrations of a system function or capability.

## **3. Stability over Time and Link to Key “Decision Dimensions” for Purchasers**

With the exception of some hospital-oriented measure sets on the JCAHO list (e.g., APACHE III, Maryland Hospital Association), none of the measure sets has been in use long enough to make a judgment about the stability over time of either individual measures in a single organization or differences among organizations. This is a key dimension of performance for any measure or indicator set being used for consumer choice purposes. The relative rankings of organizations, and the magnitude of differences among organizations, must at least be reasonably constant over periods of 3 to 5 years if the data are to have any value to consumers. Corporate purchasers and individual consumers typically make their major healthcare choices during annual open enrollment periods, generally using “report card” data at least one year old to make decisions about healthcare that will occur in the future. Individual consumers may choose a hospital or provider network on a more immediate basis if the need arises and they have an option, but the data available are typically a year or two old at best. The comparative data are only meaningful in these cases if the high-ranking and low-ranking options are the same through the period of original data collection, choice, and actual use. There is some evidence that even in the most carefully designed comparative reporting systems this condition does not hold.<sup>73</sup>

## V. Summary Evaluation of Measurement Systems

See Figure 11: Summary of Performance Measure Sets and Evaluation Criteria

**Figure 11**

<b>Summary of Performance Measure Sets and Evaluation Criteria</b>											
<b>Measure Set</b>	<b>Evaluation Criteria</b>										
	A	A1	A2	A3	B	C	D	E	F1	F2	F3
<i>Self-Assessment Instruments</i>											
SASI	Yes	?	No	No	?	Yes	?	No	?	No	?
HSIS	Yes?	?	No	No	Yes?	Yes	Yes?	No	No?	No	?
Baldrige	Yes?	Yes?	No	No	No	Yes	No	No	?	No?	?
<i>Performance Measures or Measure Sets</i>											
CRISP	Yes?	Yes?	?	?	Yes	No	Yes?	Yes	?	Most	?
FACCT	Yes	?	?	Yes	Yes	No	?	Yes?	?	Yes?	?
HEDIS	Yes	Yes?	Yes	Yes	Yes	?	Yes?	No	Yes?	Yes?	?
JCAHO	Yes?	Yes?	?	Yes	Yes	?	Yes?	Some	?	Yes?	?



## VI. Targets for the Future

### A. Vision of Performance Measurement

Having reviewed the current state-of-the-art in the area of performance measurement, it may now be possible to see a little more clearly what the future could or should look like. This field will certainly continue to evolve, but that evolution may or may not serve the specific needs of chronic care networks very well. We have already seen how some of the prominent existing measurement systems do not work well for chronic care networks, perhaps because they explicitly exclude elderly or chronically ill patients from some denominator populations or because they have chosen measures more suitable to the working-age members of private insurance plans (and their children).

From the perspective of chronic care networks, then, what sort of future performance measurement system would have “the right stuff”?

#### 1. New Measures or Sets of Measures

It is axiomatic that a measurement system for chronic care networks should include measures of chronic care. In more concrete terms, a measurement system should include measures of coordination of care among multiple providers, “seamlessness” from the patient and family point of view, costs, outcomes, and satisfaction measures from the perspective of extended episodes of care or periods of time rather than single encounters, functional health status and ability to perform activities of daily living (ADLs), caregiver burden, and the prevention of acute exacerbations or serious declines in health status. Much of this territory is new; coordination measures, for example, are essentially nonexistent in most current measurement systems because the concept is hard to operationalize.

Future measure sets for chronic care networks will probably not include measures of general preventive services or, if they do, they will use measures derived from guidelines tailored for the needs of the chronically ill. Patients with congestive heart failure do not necessarily need mammography, PSA tests, or flu shots at the same frequency as do healthy enrollees of HMOs. Some services may be unnecessary or not cost-effective; others may be absolutely crucial and must be offered on a much more frequent schedule.

#### 2. New Understanding of Process/Outcome Relationships

Measure sets that include a strong emphasis on process measures (e.g., HEDIS) are designed that way because of accepted relationships from the published literature between services provided (process) and health outcomes achieved. In the chronic care arena, the same general principle will hold, but the relationships themselves will be different. Mammography has proven benefit for healthy women in their 50s and 60s; the benefit is less or nonexistent in women with shorter life expectancy and/or more serious health risks from other diseases. Gathering information on process and outcomes in a chronic care population will add to our understanding of these relationships over time, but the measure sets of the future will be based, at least to some extent, on guidelines developed specifically for those populations.

#### 3. New Data Collection Technologies

Many of the current measurement systems, particularly those on the JCAHO list of approved measure sets for hospitals, rely primarily on administrative data (e.g., billing databases, discharge summaries) as the essential raw material. Patient surveys have a place in some measurement systems, but very few are built around manual review/abstraction of medical records, direct provision of data by physicians or other providers, or direct observation of patients. The latter technologies are difficult, expensive, and subject to a host of data quality concerns. At the same time, though, the administrative data sets typically do not include the kind of clinical detail that is necessary to either conduct case-mix or severity adjustment or focus attention on carefully defined subsets of a clinical population.

Two trends that are already apparent are likely to change this picture somewhat in the next few years. First, patients, as part of their regular interaction with providers, will be expected to provide more formal information on health outcomes and satisfaction. The format may not necessarily be a mailed survey; new telephone survey methods like Computer-Aided Telephone Interview (CATI)<sup>74</sup> or patient-initiated 800-number calls,<sup>75</sup> or even internet-based methods may make the follow-up survey process both more convenient and less expensive. Second, the expanding use of electronic medical records systems, and the linkage of those systems across care sites and provider groups, may make it easier to identify specific patient subgroups and work with existing data (rich in clinical detail) to create performance reports. The design of data structures in these systems will determine how much benefit is actually achieved; systems designed for efficient retrieval of information at the individual patient level do not necessarily do well with analyses at the aggregate or population level.<sup>76</sup>

#### **4. New Databases**

In a future in which chronic care patients have relatively long-standing relationships with a defined chronic care network, that network will have an extensive database that combines clinical, demographic, patient-reported health status, and financial information in a structure that is designed to support aggregate analysis and reporting. The individual patient (and by extension, groups of similar patients) will be the fundamental “record” in the database, rather than the individual clinical encounter. Each patient will have a single identifier for use throughout the network, so that services provided anywhere are automatically linked to the right record as soon as they are entered into the database. The database will be designed fundamentally for “population management” purposes rather than billing purposes. In order to have the same kind of coding accuracy that we currently expect in billing databases, clinicians in the network will have to adopt some level of standardization in their defining and recording of events. For example, some more explicit categorization of health status change may have to replace a notation of “patient improved” in the medical record in order to make the database maximally useful for comparative performance reports.

If patients do not link tightly to a defined chronic care network, the same sort of database may be in place, but it will not necessarily be owned and supported by the network itself. A health plan or some sort of local or regional public authority may be the owner of the database, along the lines of the “CHIN” (Community Health Information Network) model that has been promoted by the Hartford Foundation. Individual clinicians and networks will have the ability to both get information from and send information to the database for purposes of patient management. The owners or sponsors of the database will have the ability to compare the performance of networks and individual providers, subject to confidentiality constraints and sample size issues.

#### **5. New Organizational Forms**

The current measure sets are built for an organizational environment with five distinct measurement levels: individual provider, network, hospital, health plan, and integrated health system. The actual environment is more complicated than that, with many hybrids and variations, and rapid change of form of the same organization through merger, affiliation, or divestiture. The future in this area is hard to see clearly.

One trend that will affect the shape of future performance measurement systems is toward “virtual integration” rather than “vertical integration.”<sup>77</sup> Networks and systems will not necessarily be defined by clear lines of ownership and management control, but rather by more flexible relationships of contract, alliance, and “value partnership.”<sup>78</sup> The contracts, responsibility for a defined population, and shared information serve as the “glue” that holds virtual systems together. Measuring the performance of virtually integrated systems and networks requires great flexibility and ability to stay focused on core dimensions of performance while the surface structure of the organization changes rapidly.

Medical groups, hospitals, and health plans may all be relatively more stable as units of measurement and, therefore, more likely to be the focus of future measurement efforts. If a

particular multispecialty group practice can demonstrate superior performance in chronic care, that performance may be a reliable guide for consumer choice regardless of changing relationships between the group and hospitals, plans, or larger systems of care.

The ultimate question will be “who owns quality?”. The organizational form that has the strongest direct responsibility for quality, and the organizational form with the greatest ability to manage and improve quality should be the focus of performance measurement efforts. If managed care plans are performing these functions, then performance measures at the level of the health plan will make sense. On the other hand, if provider groups or networks take responsibility for quality and differentiate themselves from other networks in the area of chronic care, then performance measures at the network level will make sense. Choices about health plans will have to do primarily with which plans offer access to the best networks.

## 6. New Financing Arrangements

The current trends in healthcare financing seem to be in the direction of greater financial risk for providers in the form of capitated or “bundled” payment models. There are active research efforts underway to develop risk-adjusted capitation models that would allow plans to serve the needs of chronic care patients without incurring predictable losses, but these models do not yet have enough explanatory power to be widely accepted and used.<sup>79,80</sup> Fee-for-service models are certainly not dead yet, so performance measurement systems in the foreseeable future will have to anticipate a mixed payment environment in which some services are more likely than others to be part of a risk payment model, and the balance or level of risk may vary from region to region or plan to plan.

One implication of this “mixed model” future scenario is that performance measurement systems will have to be designed to detect both over and underservice. HEDIS was designed primarily to detect underservice (particularly in preventive or routine services) because of fear that capitation payments would provide powerful incentives to do too little. Measure sets that include measures like cesarean section rates, carotid endarterectomy rates, or proportions of biopsies with “negative” findings are designed to detect overservice because of concern that fee-for-service incentives are leading to inappropriate or unnecessary procedures. Ideally, a measure set would identify performance at the “right” level, and be able to detect departures in either direction. In practice, achieving the ideal is not a trivial matter, since measures will be designed differently depending on whether one is primarily interested in finding too much or too little of something.

## B. Where Are We Relative to This Vision?

### 1. Measures and Measure Sets

Some conceptual and “nuts and bolts” technical work needs to be done in the following areas:

- *Measuring Coordination of Care.* Starfield and colleagues at Johns Hopkins did some of the seminal work in this area in the late 1970s,<sup>81-83</sup> but there has not been significant progress since then, except perhaps for Heather Palmer’s work in both the DEMPAQ<sup>84</sup> and PROSPER<sup>85</sup> studies. Coordination is obviously a key issue for patients with serious health needs who see multiple clinicians in different physical locations, perhaps with different medical records systems. Measuring that coordination, though, and comparing levels of coordination across networks is a difficult challenge.
- *Defining Episodes of Care.* Chronic care cannot be viewed from the perspective of individual clinic encounters, specific procedures, or inpatient “episodes.” The episodes of care that matter are those that involve the onset, treatment, and possible resolution of acute

problems within the context of chronic illness, again typically involving multiple providers and sites of care. The chronic illness itself does not lend itself to measurement by episodes, but we need to decide what time periods are best suited for comparative performance measurement.

- *Measuring Costs, Outcomes, and Satisfaction.* Cost measures continue to be difficult since there is inevitably room for local variation in methods of allocating costs to departments and services. If outcomes are defined in terms of functional health status, ADL status, or freedom from pain or other symptoms, the measurement technology is rich and improving steadily. There will be a need to invest in shorter, more focused measures in order to reduce respondent burden if formal outcomes measurement becomes more routine. Satisfaction measures for individual encounters are plentiful, but there is still some work to be done on episode-of-care-oriented satisfaction measures, particularly those tailored for chronic care populations. Work at the Picker Institute is a clear step in the right direction.<sup>86</sup>
- *Case Mix Adjustment.* The existing inpatient-oriented severity of illness systems are not up to the task of case mix adjustment for chronic care populations who are primarily seen in outpatient or home settings. Greenfield's work on severity models for diabetes and other chronic conditions<sup>46</sup> is an example of the kind of work that needs to be done in order to be able to compare chronic care networks serving different populations. Work by Hornbrook,<sup>79</sup> Fowles,<sup>80</sup> and others on development of risk-adjusted capitation models will probably identify variables useful for building case mix models in the context of performance measurement.
- *Measures of Preventive Services.* Measures in this area will have to be linked to guidelines disseminated by specialty societies, AHCPR, or the NCCC. What should a chronic care network be doing about mammograms, flu shots, blood pressure monitoring, or colon cancer screening for a 70-year-old woman with congestive heart failure and diabetes? The measures are relatively easy to calculate and report once the guidelines are clear.

## **2. Process/Outcome Relationships**

This is another area where the fundamental work to be done is not about measures or measure sets per se, but about the basic science of relationships between treatment processes and outcomes in the chronic care population. Defining the relationships is somewhat more difficult here than in other clinical populations, because there is some ambiguity about interpretation of outcome measures. In oncology, one can focus on five-year relapse-free survival as an outcome, evaluate treatments against that outcome, and eventually derive guidelines and performance measures based on what is known about the relationship between drugs, doses, timing of treatments, and five-year survival. For chronic conditions with significant morbidity associated, the outcomes may be survival-related, but they may also be defined in terms of functional independence, freedom of pain, ADL level, amount of caregiver burden, or avoidance of specific complications or disease exacerbations. Different treatment processes will be more clearly linked to some outcomes than others. Advances in "clinical epidemiology" linking processes to outcomes for chronic care patients will probably be the most important factor underlying advances in performance measurement.

## **3. Data Collection Technologies**

The DEMPAQ and PROSPER projects have demonstrated the feasibility of using medical records review and patient surveys, respectively, to measure quality of care in ambulatory settings. Some quality improvement projects carried out by Physician Review Organizations (PROs) have shown the same thing. In spite of these successes in individual research projects, though, the technology for medical record review remains cumbersome and expensive. We do not yet have data systems capable of integrating key clinical variables (e.g., test results, lab

values), patient demographics, patients' self-reported functional status, and clinician ratings in a way that permits regular, ongoing, "real-time" calculation of performance measures. Some of the best electronic medical records systems are beginning to approach this kind of capability,<sup>87</sup> but most chronic care networks are a long way away from being able to do this.

Our techniques for administering patient health status surveys are still typically based on paper-and-pencil methods through either in-clinic administration or mail and return. There is a need for testing of better, more convenient methods. One interesting new approach is the use of automated 800-number lines that patients can call at any time and work through a set of pre-recorded questions using a touch-tone phone to provide answers.<sup>75</sup> Another interesting approach is the use of item response theory-based computerized adaptive testing (IRT-based CAT) branching techniques that use the minimum number of questions to assign a score or category to a patient on a particular health status or satisfaction dimension.<sup>88</sup> In general, we need to develop survey and interview methods that are shorter, less intrusive, and more flexible for patients in terms of when they provide the information and how.

#### **4. New Databases**

Some integrated health systems and chronic care networks are making exciting progress in the direction of comprehensive patient databases suitable for use in performance measurement. Group Health Cooperative of Puget Sound has been a leader in this area for some time; Henry Ford Health System and the Sisters of Providence are two other examples of organizations that are actively working to develop population-oriented databases that permit analysis of care patterns and outcomes for chronic care patients.

#### **5. Organizational Forms and Performance Databases**

We still have a great deal of work to do to align the "network" organizational form for chronic care with the organizational units that maintain and use performance measurement systems. Health plans using HEDIS 2.5 or 3.0, for example, will have only a few measures that apply to chronic care (see pages 17-18), even in situations where the health plan is functioning in the role of "network manager." (NCQA is working on a set of HEDIS measures for chronic disease with the support of the Robert Wood Johnson Foundation. Work groups have done much of the background work and have prepared reports, but the measures have not yet been pilot-tested by NCQA or included in any HEDIS version, including the testing set measures in HEDIS 3.0.<sup>89</sup>) The CRISP measures are designed to apply to networks/organized delivery systems, but again, the current measures are not focused on chronic care. The JCAHO-approved measure sets have mainly to do with hospital care; the five measure sets for networks and health plans include HEDIS and FACCT, then three specific sets for nursing homes, home healthcare, and rehabilitation facilities. None of these measure sets are ideally suited for the needs of chronic care networks. It may be useful to select individual measures from these sets that are appropriate as a starting framework, then do some consensus building and pilot testing to fill in the gaps.

#### **6. Match to Current Financing Arrangements**

As mentioned earlier, HEDIS is designed primarily to fit a world of managed care, capitation payments, and potential underservice to plan members by plans or providers motivated to control costs. The other measurement systems or self-assessment programs are not designed as clearly around a payment model and related incentives. Given the uncertainty about the future payment environment for chronic care, this may be a good situation. If the measure sets and individual measures within them can continue to evolve in the direction of "ideal ranges" or "standards of excellent care" and then have the ability to detect departures from optimal in either direction, the measures should be robust to any future changes in financing arrangements.

## **C. Moving from Current State to Future Vision—How Do We Get There?**

The currently available measure sets have come into being through several mechanisms: sponsorship of a national organization (SASI by NCCC, HEDIS by NCQA and the HMO Group, the ORYX system by JCAHO); extensive purchaser input (HEDIS, FACCT); technical input from recognized academic and clinical expert consultants (HEDIS, FACCT, CRISP); interest by single investigators, single organizations, or small voluntary groups of organizations (HSIS, CRISP), or official government sponsorship (Malcolm Baldrige criteria). All of these factors will probably continue to play a role in the continuing evolution of performance measurement systems as well.

### **1. Refining Existing Measures or Developing New Measures**

HEDIS, JCAHO, and CRISP all have mechanisms in place for testing, refining, and incorporating new measures into existing measure sets. Feedback on existing measures is used to refine or modify them. These mechanisms include explicit feedback by users, presence of technical work groups or expert panels, and a commitment to “continuous quality improvement” of the measure sets themselves. HEDIS, CRISP, and FACCT have been active in seeking external grant support from private foundation and government sources for some of this refinement or new measures development work.

The National Chronic Care Consortium and its member organizations can be an effective voice for suggesting (even demanding) the inclusion of relevant measures in these measure sets. Formal participation in the projects, service on technical advisory committees, and regular communication with project staff all are available (and often welcome) means of influencing the course of future developments.

Purchasers have, and will continue to have, the last say in the design of performance measurement systems, at least those designed for external reporting/comparative purchasing purposes. The HEDIS and FACCT measure sets are in place and constructed the way they are largely due to the information demands of major healthcare purchasers. Any purchasers who focus attention specifically on chronic care networks and their performance will be in a strong position to define the measure sets that best meet their needs (and those of their employees or constituents). If purchasers act independently, the result will probably be a very confusing and difficult situation in which each purchaser and each contract is associated with a slightly (or very) different set of performance measurement requirements. Purchaser coalitions in a local market area or on a national scale (FACCT) can serve a crucial “consensus-building” function in minimizing the number of competing data requests. An accrediting body (e.g., JCAHO, NCQA) can perform the function even more efficiently, simply by making a single measure set required. If chronic care networks become subject to accreditation requirements either through the JCAHO or NCQA, a standard set (or accepted list of sets) of performance measures will almost certainly be part of the accreditation process.

### **2. Building an Understanding of Process/Outcome Relationships**

The key activity here will be clinical research and/or analysis of existing research findings, rather than measures of development per se. Performance measures can be built on knowledge about process-outcome relationships, but they typically do not identify those relationships themselves. One exception to this general pattern is in the area of process-outcome relationships at the organization level. We generally presume that if processes relate to outcomes strongly enough at the single patient level, that organizations with better processes should have better outcomes. It is not well understood, though, what size difference in process at the organization level is associated with meaningful outcome differences at that level. For example, how much better must one organization be at monitoring and controlling hypertension in order for a prospective new patient to expect a significant reduction in the likelihood of stroke? The answer depends not only on the biology of hypertension control and stroke, but on the statistical issues related to extrapolating from aggregate data to the individual case and the stability of organizational performance in the process realm over relatively long periods of time. If clear, statistically significant excellence

in hypertension control is associated with only a one percent absolute reduction in the likelihood of stroke over the next ten years, is it worth switching plans or networks and perhaps paying more in premiums to gain that benefit? What if the network is good at hypertension control but not good at glycemic control for diabetes? How does a patient with both hypertension and diabetes translate that process information into something meaningful about likely health outcomes for him or her? Measurement efforts with large numbers of organizational members may be able to begin to address this issue by analyzing the relationship between process differences among plans or networks at one time period at outcome differences in later time periods. Some of the published work on the relationship between hospital quality characteristics and mortality rates is a model for this kind of study.<sup>90</sup>

### **3. Developing New Data Collection Technologies**

Many interesting technologies are already on the horizon and just need further testing and development. Some of this work can be supported in the private sector, as companies seek to develop commercial capabilities that can be applied in a large number of survey or interview applied areas.<sup>75</sup> The Agency for Health Care Policy and Research can consider funding of “data technology” studies as part of its larger agenda on quality of care measurement. Voluntary associations like NCCC, the Group Practice Improvement Network (GPIN), or the Institute for Healthcare Improvement (IHI) can play a crucial information exchange role as members share “best practices” with each other in this area.

### **4. Creating New Databases**

Most health plans or delivery networks are already making extensive investments in information systems<sup>91</sup>; the key question may be how those efforts can be channeled to produce the kind of capabilities needed for performance management in the future. Some of the channeling comes directly from external pressures. Health plans design systems to collect and report HEDIS data because such reporting is required of them. Billing databases are built around the fields included in the UB-92 or the HCFA-1500 form because those forms and their data elements are required by purchasers.

There is a less clear mandate to build databases around the characteristics of populations. The CHIN “movement” is proceeding slowly, if at all, because there is no clear consensus about the need for, or value of, community-level healthcare databases.<sup>92</sup> Organizations with a clear mission statement about population health may invest in clinical databases with that purpose in mind, but the internal competition for resources is fierce when there are unmet needs in areas like clinic scheduling, billing, and laboratory results reporting. Again, purchasers and the individuals they represent have the final say. If plans and networks are required to report data on quality of care to defined populations, both in terms of processes, outcomes, and cost-effectiveness, the databases to meet that requirement will be created. The cost will be borne by the ultimate consumer, the individual patient, so there must be a clear sense in the community that such databases are worth the investment.

### **5. Creating New Organizational Structures and Developing Related Measurement Structures**

“Virtual integration” may be perhaps the most useful concept to keep in mind when thinking about how to develop truly integrated systems of care for chronically ill patients. It is not strictly necessary that a single health plan, group practice, or integrated delivery system own and manage all parts of the care continuum in order for that continuum to function effectively. Contractual relationships and alliances can accomplish the integration and coordination as well, as long as there are mechanisms in place for:

- creating, implementing, and enforcing clinical guidelines and pathways;
- allocating capital and other sorts of resource investments to those areas most likely to produce benefits in terms of patient care;

- supporting a systemwide quality measurement and continuous quality improvement program;
- creating and nurturing a common culture in areas like customer service, sensitivity to patient/family preferences, mutual respect and tolerance, and ethical business practices;
- assuring access to the complete spectrum of clinical and support services, including those best provided by community-based agencies;
- aligning the possible competing incentives caused by different payment systems, and resolving conflicts in favor of the patients' best interests.

There seems to be no good way to predict or control the development of organizational forms in the future. There are no natural, simple scientific laws that govern this development; "chaos theory" may be the most applicable scientific body of thought. Prevailing wisdom changes regularly. Flexibility may be the most important organizational attribute.

In this chaotic environment, measurement systems cannot be too closely linked to organizational forms that may go out of fashion overnight. A measurement system designed for health plans with a relatively fixed denominator population of "members" may find itself out of date and unsuitable in an environment of "carve-outs" and direct contracting between purchasers and provider networks. Likewise, a measurement system organized around particular facilities like hospitals or nursing homes may be left behind if the treatment environment moves to home or community settings. The best advice may be to keep the measurement systems oriented to populations of chronic care patients and their needs, and maintain maximum flexibility about the organization from which the population receives care.

## **6. New Financial Arrangements**

Capitation payments or other forms of bundled payment that involve large parts of the care continuum serve as a driving force for network formation and network performance measurement. Fee-for-service payment, on the other hand, provides no particular motivation for network formation and may tend to keep data systems (e.g., billing databases) separate from each other. The single most important advance in this area would be a risk-adjusted capitation model designed for chronic care patients. If a coalition of purchasers, providers, and government agencies could reach some working consensus about a model and implement it at least on a demonstration project basis, the stage would be set for a "defined population" approach to performance measurement.

In the meantime, some more focused carve-out payment models may be easier to implement. Care of ESRD or AIDS patients, for example, is already running on a separate payment method, at least in some organizations. The PACE/On Lok program<sup>93</sup> is an example of bundled payment for a specific class of chronic care patients (frail elderly). Other examples can be found in state Medicaid programs where chronic mental illness is funded and provided through separate carve-out contracts with community mental health boards.

## VII. Conclusions

### A. How Good Is the Current “State-of-the-Art”?

There are three excellent, detailed tools available for organizational self-assessment by chronic care networks. The SASI tool is specifically designed for that purpose; the HSI and Malcolm Baldrige criteria are more generic but certainly applicable. None of the performance measure sets reviewed in this report—HEDIS, FACCT, CRISP, JCAHO—are designed for chronic care networks. Individual measures within those larger sets are applicable and useful, but the sets cannot be taken as a whole in their current form and used to describe or compare the performance of chronic care networks. The measure sets are good, but in domains other than chronic care.

### B. What Needs to Be Done?

Using the existing measure sets as a starting point, a measure set specific to the care processes and significant outcomes for chronic care populations needs to be developed. Development of the measure set can precede development of all of the required data infrastructure, since the need to gather data for measurement can serve as the key motivator and design criterion for new data systems. The self-assessment tools can serve right now as a vehicle for examining chronic care networks’ structures, internal relationships, and capabilities; a performance measurement system can illuminate what the networks actually are able to do.

The essential pieces are all in place to move forward. The provider networks already exist, and have national organizations (NCCC, possibly along with GPIN, the HMO Research Network, and the AMGA) to serve as a focal point for collaborative efforts. Purchasers have established a role in the design and use of quality measurement systems, and have their own organizations as well (Washington Business Group on Health, Pacific Business Group on Health, FACCT). Purchasers just need to identify chronic care and chronic care networks as entities worth attending to in the same way that health plans and hospitals are currently. The federal government, through HCFA, has mechanisms for identifying chronic care as a priority area, defining new payment models, and conducting demonstration projects. AHCPR can support some of the academic research needed to move forward in the technology of measurement. State governments also have a contribution to make through the inclusion of chronic care patient populations in Medicaid.

If these organizations are all pulling in different directions, we will see an extended period of confusion and false starts. If, on the other hand, there is general consensus about the direction in which chronic care and its measurement need to move, there will be plenty of room for innovation and creativity about the details.



---

## References

1. Griffith, JR. "Managing the Transition to Integrated Health Care Organizations." *Frontiers of Health Services Management* (1996) 12:4-50.
2. Bushick, JB. "Creating an Integrated Health Care System as a Basis for Managed Care Excellence: A Health Plan's Perspective." *Managed Care Quarterly* (1995) 3:1-10.
3. Conrad, DA and Dowling, WL. "Vertical Integration in Health Services: Theory and Managerial Implications." *Health Care Management Review* (1990) 15:9-22.
4. Warden, GW, Zajac, BM, and Nerenz, DR "Vertically Integrated Health Systems: An Option for Health Care Reform." *Decisions in Imaging Economics* (1991) 4:27-31.
5. Lansky, D and Purdy, S. "Public Accountability for Health: New Standards for Health System Performance." *Managed Care Quarterly* (1995) 3:17-24.
6. Ross, A and Fenster, LF. "The Dilemma of Managing Value." *Frontiers of Health Services Management* (1995) 12:3-32.
7. Wennberg, JE and Gittelsohn, A. "Small Area Variations in Health Care Delivery." *Science* (1973) 182:1102-1108.
8. Wennberg, JE and Gittelsohn, A. "Variations in Medical Care Among Small Areas." *Scientific American* (1982) 246:120-133.
9. Wennberg, JE. "Dealing with Medical Practice Variation: A Proposal for Action." *Health Affairs* (1984) 3:7-31.
10. Bernstein, SJ; McGlynn, EA; Siu, AL; et al. "The Appropriateness of Hysterectomy: A Comparison of Care in Seven Health Plans." *JAMA* (1993) 269:2398-2402.
11. Park, RE, Fink, A, Brook, RH, et al. "Physician Ratings of Appropriate Indications for Six Medical and Surgical Procedures." *American Journal of Public Health* (1986) 76:766-772.
12. Ellwood, PM. "Shattuck Lecture—Outcomes Management: A Technology of Patient Experience." *New England Journal of Medicine* (1988) 318:1549-1556.
13. Marwick, C. "New Health Care Research Agency Reflects Interest in Evaluating Quality." *JAMA* (1990) 263:929-930.
14. National Committee for Quality Assurance. *Health Plan Employer Data and Information Set and Users' Manual (Version 2.0)*. Washington, DC: NCQA (1993).
15. Wagner, EH. "Managed Care and Chronic Illness: Health Services Research Needs." *Health Services Research* (1997) 32:702-714.
16. Von Korff, M; Gruman, J; Schaefer, S; et al. "Collaborative Management of Chronic Illness: Essential Elements." *Annals of Internal Medicine* (1997).
17. Wagner, EH; Austin, BT; VonKorff, M. "Organizing Care for Patients with Chronic Illness." *Milbank Quarterly* (1996) 74:511-544.

18. National Chronic Care Consortium. "NCCC Member Current Practices." National Chronic Care Consortium. Bloomington, MN (1997).
19. Center for Health Care Policy and Evaluation. "United HealthCare Corporation Performance Indicators for Evaluating a Health Care Delivery System: A Report Card on Health." United HealthCare Corporation. Minneapolis, MN (1993).
20. Kaiser Permanente. "1993 Quality Report Card." Kaiser Foundation Health Plans. Oakland, CA (1994).
21. Rosenthal, GE and Harper, DL. "Cleveland Health Quality Choice: A Model for Collaborative Community-Based Outcomes Assessment." *Joint Commission Journal on Quality Improvement* (1994) 20:425-442.
22. Lipson, DJ and De Sa, JM. "Impact of Purchasing Strategies on Local Health Care Systems." *Health Affairs* (1996) 15:62-76.
23. Corrigan, JM. "How Do Purchasers Develop and Use Performance Measures?" *Medical Care* (1995) 33:JS18-JS24.
24. Corrigan, JM and Nielsen, DM. "Toward the Development of Uniform Reporting Standards for Managed Care Organizations—The Health Plan Employer Data and Information Set (Version 2.0)." *Joint Commission Journal on Quality Improvement* (1993) 19:566-575.
25. National Committee for Quality Assurance. "HEDIS 3.0." National Committee for Quality Assurance. Washington, DC: (1996).
26. ORYX PLUS: Specifications for Performance Measures. Joint Commission on Accreditation of Healthcare Organizations. Oakbrook Terrace, IL (1997).
27. Bethell, C and Lansky, D. "Foundation for Accountability: Framework for accountability." In Seltzer, J and Nash, DB (Eds.) "Models for Measuring Quality in Managed Care: Analysis and Impact." Faulkner & Gray (1997) 55-78.
28. Nerenz, DR and Zajac, BM. "Indicators of Performance for Vertically Integrated Health Systems." American Hospital Association. Chicago, IL (1991).
29. Nerenz, DR; Zajac, BM; Rosman, HS. "Consortium Research on Indicators of System Performance (CRISP)." *Joint Commission Journal on Quality Improvement* (1993) 19:577-585.
30. Zajac, BM; Green-Weir, R; Nerenz, DR. "Health Status of Populations as a Measure of Health System Performance." *Managed Care Quarterly* (1995) 3:29-38.
31. Shortell, SM; Gillies, RR; Anderson, DA; et al. "Remaking Health Care in America: Building Organized Delivery Systems." Josey-Bass Publishers (1996).
32. Shortell, SM; Gillies, RR; Anderson, DA; et al. "Creating Organized Delivery Systems: The Barriers and Facilitators." *Hospital and Health Services Administration* (1993) 38:447-466.
33. National Chronic Care Consortium. "SASI: Self-Assessment for Systems Integration Tool." National Chronic Care Consortium. Bloomington, MN (1995).

34. Berwick, DM. "Continuous Improvement as an Ideal in Healthcare." *New England Journal of Medicine* (1989) 320:53-56.
35. Batalden, PB and Stoltz, PK. "A framework for the Continual Improvement of Healthcare: Building and Applying Professional and Improvement Knowledge to Test Changes in Daily Work." *Joint Commission Journal on Quality Improvement* (1993) 19:424-447.
36. Mohr, JJ; Mahoney, CC; Nelson, EC; et al. "Improving Healthcare, Part 3: Clinical Benchmarking for Best Patient Care." *Joint Commission Journal on Quality Improvement* (1996) 22:599-616.
37. Berwick, DM. "Harvesting Knowledge From Improvement." *JAMA* (1996) 275:877-878.
38. Laffel, G and Blumenthal, D. "The Case for Using Industrial Quality Management Science in Healthcare Organizations." *JAMA* (1989) 262:2869-2873.
39. U.S. Department of Commerce. "Malcolm Baldrige National Quality Award: Criteria for Performance Excellence." Commerce Department. Washington, DC (1997).
40. Knaus, WA; Wagner, DP; Draper, EA; et al. "The APACHE III Prognostic System: Risk Prediction of Hospital Mortality for Critically Ill Hospitalized Adults." *Chest*. (1991) 100:1619-1636.
41. Hibbard, JH and Jewett, JJ. "Will Quality Report Cards Help Consumers?" *Health Affairs* (1997) 16:218-228.
42. Hibbard, JH and Jewett, JJ. "What Type of Quality Information Do Consumers Want in a Healthcare Report Card?" *Medical Care Research Review* (1996) 53:28-47.
43. Solberg, LI; Mosser, G; McDonald, S. "The Three Faces of Performance Measurement: Improvement, Accountability, and Research." *Joint Commission Journal on Quality Improvement*. (1997) 23:135-147.
44. Foundation for Accountability (FACCT). "Guidebook for Performance Measurement: Prototype." Foundation for Accountability, Portland, OR (1995).
45. Iezzoni, LI. (Ed.) "Risk Adjustment for Measuring Health Care Outcomes." Health Administration Press, Ann Arbor, MI (1994).
46. Greenfield, S; Sullivan, L; Dukes, KA; et al. "Development and Testing of a New Measure of Case Mix for Use in Office Practice." *Medical Care* (1995) 33:AS47-AS55.
47. Weiner, J; Starfield, B; Steinwachs, D; et al. "Development and Application of a Population-Oriented Measure of Ambulatory Care Case-Mix." *Medical Care* (1991) 29:452-472.
48. Hibbard, JH; Sofaer, S; Jewett, JJ. "Condition-Specific Performance Information: Assessing Salience, Comprehension, and Approaches for Communicating Quality." *Health Care Financing Review* (1996) 18:95-109.
49. Epstein, A. "Performance Reports on Quality—Prototypes, Problems, and Prospects." *New England Journal of Medicine* (1995) 333:57-61.

50. Tanenbaum, SJ. Knowing and Acting in Medical Practice: The Epistemological Politics of Outcomes Research. *Journal of Health Politics, Policy, and Law*, 1994, 19, 27-44.
51. National Committee for Quality Assurance. "Health Plan Employer Data and Information Set and Users' Manual (Version 2.0)." National Committee for Quality Assurance. Washington, DC (1993).
52. Roberts, J and Corrigan, JM. "Commentaries on Public Accountability." *Managed Care Quarterly* (1995) 3:25-29.
53. Gosfield, A. "Who is Holding Whom Accountable for Quality?" *Health Affairs* (1997) 16:26-40.
54. Pacific Business Group on Health. 1996 Report on PBGH Quality Initiatives. San Francisco, CA: Pacific Business Group on Health, 1996.
55. Gillies, RR; Shortell, SM; Anderson, DA; et al. "Conceptualizing and Measuring Integration: Findings from the Health Systems Integration Study." *Hospital and Health Services Administration* (1993) 38:467-489.
56. Shortell, SM; Gillies, RR; Anderson, DA. "The New World of Managed Care: Creating Organized Delivery Systems." *Health Affairs* (1994) 13:46-64.
57. Devers, KJ; Shortell, SM; Gillies, RR; et al. "Implementing Organized Delivery Systems: An Integration Scorecard." *Health Care Management Review* (1994) 19:7-20.
58. Hertz, HS; Reimann, CW; Bostwick, MC. "The Malcolm Baldrige National Quality Award Concept: Could it Help Stimulate or Accelerate Healthcare Quality Improvement?" *Quality Management in Health Care* (1994) 2:63-72.
59. Berman, S. "Putting it (Performance Measures) Together: An Interview With Jerod Loeb." *Joint Commission Journal on Quality Improvement* (1996) 22:518-526.
60. "ORYX: Technical Implementation Guide for Performance Measurement Systems." Joint Commission on Accreditation of Healthcare Organizations. Oakbrook Terrace, IL. (1997).
61. "Comprehensive Accreditation Manual for Health Care Networks." Joint Commission on Accreditation of Healthcare Organizations. Oakbrook Terrace, IL. (1996).
62. "APACHE Acute Care Series." APACHE Medical Systems, Inc. McLean, VA.
63. Kazandjian, VA; Wood, P; Lawthers, J. "Balancing Science and Practice in Indicator Development: The Maryland Hospital Association Quality Indicator Project." *International Journal for Quality in Health Care* (1995) 7:39-46.
64. Robinson, S and Brodie, M. "Understanding the Quality Challenge for Health Consumers: The Kaiser / AHCPR Survey." *Joint Commission Journal on Quality Improvement* (1997) 23:239-244.
65. Sofaer, S. "How Will We Know if We Got it Right? Aims, Benefits, and Risks of Consumer Information Initiatives." *Joint Commission Journal on Quality Improvement* (1997) 23:258-264.

66. Allen, HM and Rogers, WH. "The Consumer Health Plan Value Survey: Round Two." *Health Affairs* (1997) 16:156-166.
67. United States General Accounting Office. "'Report Cards' Are Useful but Significant Issues Need to Be Addressed." GAO/HEHS. Washington, DC (1994) 94-219.
68. Committee on Performance Measurement. "HEDIS 3.0." National Committee for Quality Assurance. Washington, DC (1997).
69. Spoeri, RK and Ullman, R. "Measuring and Reporting Managed Care Performance: Lessons Learned and New Initiatives." *Annals of Internal Medicine* (1997) 127:726-732.
70. Ward, RE. "HEDIS 3.0 and Beyond: The Future of Performance Measurement." *Quality Source for America's Leading Medical Groups* (1996) 4:5-9.
71. Nerenz, DR and Zajac, BM. "Assessing Performance of Integrated Delivery Systems: Analyzing Indicators and Reporting Results." Faulkner & Gray. New York (1996).
72. Berman, S. "Using Malcolm Baldrige National Quality Award Criteria for Improvement: An Interview with Ellen Gaucher." *Joint Commission Journal on Quality Improvement*, (1995) 21:249-256.
73. Green, J and Whitfield, N. "Report Cards on Cardiac Surgeons: Assessing New York State's Approach." *New England Journal of Medicine* (1995) 332:1229-1232.
74. Anie, KA; Jones, PW; Hilton, SR; Anderson, HR. "A Computer-Assisted Telephone Interview Technique for Assessment of Asthma Morbidity and Drug Use in Adult Asthma." *Journal of Clinical Epidemiology* (1996) 49(6):653-656.
75. Fisher, S; Kent, TA; Bryant, SG. "Postmarketing Surveillance By Patient Self-Monitoring: Preliminary Data for Sertraline Versus Fluoxetine." *Journal of Clinical Psychiatry* (1995) 56(7):288-296.
76. Greenlick, MR; Freeborn, DK; Pope, CR (Eds.). "Health Care Research in an HMO: Two Decades of Discovery." The Johns Hopkins University Press. Baltimore, MD (1988) Chapter 3.
77. Goldsmith, J. "The Illusive Logic of Integration." *Healthcare Forum Journal* (1994) 5:26-31.
78. Robinson, JC. "The Dynamics and Limits of Corporate Growth in Healthcare." *Health Affairs* (1996) 15:156-169.
79. Hornbrook, MC and Goodman, MJ. "Chronic Disease, Functional Health Status, and Demographics: A Multi-Dimensional Approach to Risk Adjustment." *Health Services Research* (1996) 31:283-308.
80. Van de Ven, WPMM; Van Vliet, RCJA; Van Barneveld, EM; et al. "Risk-Adjusted Capitation: Recent Experiences in the Netherlands." *Health Affairs* (1994) 13:120-136.
81. Starfield, B; Simborg, D; Johns, C; Horn, S. "Coordination of Care and its Relationship to Continuity and Medical Records." *Medical Care* (1977) 15:929-938.

82. Starfield, BH; Simborg, DW; Horn, SD; Yourtee, SA. "Continuity and Coordination in Primary Care: Their Achievement and Utility." *Medical Care* (1976) 14:625-636.
83. Fletcher, RH; O'Malley, MS; Fletcher, SW; et al. "Measuring the Continuity and Coordination of Medical Care in a System Involving Multiple Providers." *Medical Care* (1984) 22:403-411.
84. Lawthers, AG; Palmer, RH; Edwards, JE; et al. "Developing and Evaluating Performance Measures for Ambulatory Care Quality: A Preliminary Report of the DEMPAQ Project." *Joint Commission Journal on Quality Improvement* (1993) 19:552-565.
85. Zapka, JC; Palmer, RH; Hargraves, LJ; et al. "Relationship of Patient Satisfaction with Experience of System Performance and Health Status." *Journal of Ambulatory Care Management* (1995) 18:78-83.
86. Edgman-Levitan, S and Cleary, PD. "What Information Do Consumers Want and Need?" *Health Affairs* (1996) 15:42-56.
87. Detmer, DE; Steen, EB; Dick, RS. "The Computerized Patient Record: An Essential Technology for Health Care." National Academy Press. Washington, DC (1997).
88. McHorney, CA. "Generic Health Measurement: Past Accomplishments and a Measurement Paradigm for the 21st Century." *Annals of Internal Medicine* (1997) 127:743-750.
89. Kuhlthau, K; Klein Walker, D; Perrin, J; Bauman, L; et-al. "Assessing Managed Care For Children With Chronic Conditions." *Health Affairs* 17(4):42-52.
90. Hartz, AJ; Gottlieb, MS; Kuhn, EM; Rimm, AA. "The Relationship Between Adjusted Hospital Mortality and the Results of Peer Review." *Health Services Research* (1993) 27:765-777.
91. Siwicki, B. "Crafting New Methods of Systems Integration." *Health Data Management*, (1997) 5:83-88.
92. Starr, P. "Smart Technology, Stunted Policy: Developing Health Information Networks." *Health Affairs* (1997) 16:91-105.
93. Eng, C; Pedulla, J; Eleazer, G; McCann, R; Fox, N. "Program of All-Inclusive Care for the Elderly (PACE): An Innovative Model of Integrated Geriatric Care and Financing." *Journal of the American Geriatrics Society* (1997) 45:223-232.