The Practice of Informatics

JAMIA

White Paper ■

An Informatics Blueprint for Healthcare Quality Information Systems

JOYCE C. NILAND, PhD, LAYLA ROUSE, MS, DOUGLAS C. STAHL, PhD

A b s t r a c t   There is a critical gap in our nation's ability to accurately measure and manage the quality of medical care. A robust healthcare quality information system (HQIS) has the potential to address this deficiency through the capture, codification, and analysis of information about patient treatments and related outcomes. Because non-technical issues often present the greatest challenges, this paper provides an overview of these socio-technical issues in building a successful HQIS, including the human, organizational, and knowledge management (KM) perspectives. Through an extensive literature review and direct experience in building a practical HQIS (the National Comprehensive Cancer Network Outcomes Research Database system), we have formulated an "informatics blueprint" to guide the development of such systems. While the blueprint was developed to facilitate healthcare quality information collection, management, analysis, and reporting, the concepts and advice provided may be extensible to the development of other types of clinical research information systems.

Introduction

Physician practice patterns and corresponding treatment outcomes vary much more widely than previously realized, and such variations have been associated with both sub-optimal patient outcomes and increased treatment costs.[1] Observed differences in treatment outcomes across populations suggest that major opportunities for improvement exist, and clinicians, patients, and the general public are demanding more information about healthcare quality.[2–4] Healthcare consumers seek to become better informed about their choices, and expect to see provider-specific clinical outcomes data to confirm the promised benefits of medical treatments.[5] Payors require clinical outcome data to evaluate quality of care and cost-effectiveness.[6]

It remains extremely difficult to accurately measure quality indicators in the general patient population, and to relate these measurements to outcomes. A healthcare quality information system (HQIS) is a data system that captures data on medical programs and practices, and provides monitoring and reports on care and outcomes for patients treated within the caregiver system. Thus a HQIS offers the opportunity to address current deficiencies in healthcare evaluation, and to ultimately improve the quality of care.[7,8]

Since 1996 the Division of Information Sciences at City of Hope National Medical Center has been responsible for the development, implementation, and maintenance of a multi-centered Internet-based HQIS, the National Comprehensive Cancer Network (NCCN) Outcomes Research Database System.[9–11] The NCCN is a volunteer organization of 20 Cancer Centers around the country, formed to continually improve the quality of oncology care. The first objective of the NCCN was to establish a robust set of care guidelines based on clinical trials evidence and expert opinion; these guidelines now exist for over 95% of all cancer types, and are updated at least annually as new medical evidence emerges.

The next goal of the NCCN was to measure guideline concordance across participating Cancer Centers, assess patterns of care over time, and benchmark that care against quality indicators and patient outcomes. The NCCN HQIS was developed to collect standardized coded data on all patients receiving primary care at a participating NCCN institution, and to facilitate analysis of these data to support delivery of the highest possible quality of cancer care. After initial treatment at the NCCN center and at specified follow-up intervals thereafter, coded data are abstracted by Clinical Research Associates, based on existing medical records and patient surveys obtained during the routine care process. The NCCN HQIS was designed to:

- capture coded demographic, diagnosis, co-morbidity, treatment and outcomes data into a single centralized data repository located at City of Hope;
- measure concordance of the treatment given with the NCCN oncology care guidelines;
- analyze associations between patient outcomes and patient demographics, treatment factors, and Cancer Center characteristics.

Affiliation of the authors: Division of Information Sciences, City of Hope National Medical Center, Duarte, CA.

Correspondence and reprints: Joyce C. Niland, PhD, Division of Information Sciences, City of Hope National Medical Center, 1500 East Duarte Road, Duarte, CA 91010; e-mail: ⟨jniland@coh.org⟩.

Relatively few clinical research studies were utilizing the Web for data collection/transmission in 1996. However, recognizing the power and efficiencies that could be gained, our design specifications called for the implementation of a Web-enabled relational database that follows a client-server model. The Web interfaces were created using Microsoft Active Server Page (ASP) technology and JavaScript, on Microsoft Internet Information Server (IIS) and SQL Server platforms. The NCCN HQIS was constructed over a robust security framework that includes role-based data access, data encryption, and digital certification. Authorized users can enter the data directly into the central repository via Web-based screens that invoke logic checks, skip patterns, and conditional drop-down menus. Alternatively, for centers that already have existing electronic data that matches or can be mapped to the NCCN data dictionary, an electronic data file may be transmitted to the DCC on a quarterly basis, for uploading into the pooled repository.

Through the process of developing the NCCN outcomes research system, an appreciation of the many challenges and difficulties in creating a robust HQIS has been gained. While descriptions of information systems typically focus on the technology involved, a successful HQIS consists of much more than hardware and software. In fact the socio-technical considerations of the organizational culture, data structure, and knowledge management processes of the HQIS often prove to be far more challenging in building such systems.[4,7,8]

To be well accepted, any new technical system has to be part of a working socio-technical system, engaging with the complex world of tasks, procedures and culture within an organization.[12] Otherwise high failure rates and low utilization levels are certain consequences. Based on our past experience and ongoing efforts to fully optimize the NCCN outcomes research database system, we were motivated to identify and summarize the most important HQIS socio-technical characteristics as described in the informatics literature.

In this paper we first document the critical need for information systems to capture practice patterns and outcomes, to facilitate benchmarking against established healthcare quality guidelines. We then present our review of the relevant literature, organized around several higher order themes that form the foundation of an "informatics blueprint" for HQIS efforts. While this review was motivated by our NCCN HQIS to support oncology outcomes research, the resulting blueprint serves as a guide to the construction and evaluation of information systems across many related domain areas.

Background

The Demand for Outcomes Data

Outcomes research is the science of accurately measuring treatment patterns in the general patient population, and relating those patterns to patient outcomes. With the growing human and financial burdens attributed to chronic illnesses, the demand for outcomes data and healthcare quality measurement is increasing on a national scale. Stakeholders include patients, healthcare providers, purchasers, accrediting organizations, professional societies, and government agencies. This "outcomes movement" has been fueled by research that describes substantial geographical differences in hospital admissions and medical procedures, differences that cannot be explained solely by severity of illness.[13] Such practice variations are driven by many factors, including patient population differences, lack of professional consensus, non-uniform access to care, differences in local or regional capabilities, and the overall quality of care practices. The great concern is that this practice variability may lead to suboptimal treatment for a substantial proportion of patients.[1,14,15]

To facilitate healthcare quality, there is a burgeoning interest in the relationship between healthcare processes and outcomes, leading to a growing demand for data to support such studies.[13,16] Patients want to understand treatment risks and benefits, and to receive the best possible outcomes from their selected treatment regimens. Clinicians are increasingly interested in objective information regarding their own practice patterns, and the ability to benchmark their treatment practices against peers within the medical community. The pursuit of answers to these questions requires empirical data to allow assessments of practice patterns and corresponding intervention outcomes.

Healthcare quality measurement is an elusive goal, and current quality of care measurement practices are relatively primitive.[17] There is a paucity of data to assess the implementation of treatment guidelines and related treatment outcomes.[18] In an effort to monitor and improve care, insurers and managed-care groups often apply utilization review, profiling, and other rudimentary methods.[4] Such approaches are largely based on administrative or billing data, and lack the clinical details to accurately evaluate treatment outcomes while adjusting for confounding factors such as co-morbidity.[18] Limited diagnostic coding in administrative databases restricts the amount of clinical detail that can be captured, making it difficult to discern important temporal distinctions between existing co-morbidity and complications of the care itself.[19–21]

Through the implementation of a robust HQIS raw data can be transformed into useful codified information, leading to new knowledge that may improve patient care. These systems must support a particular form of knowledge management (KM), defined as "the process of creating, capturing, and using knowledge to enhance organizational performance."[22] The development of large, sophisticated databases will be required to support complex analyses involved in effective outcomes monitoring and management.[23,24]

Requirements of a Successful HQIS

In addition to the requisite hardware and software, HQIS developers must specify a complete "information framework,"[25] including the standard data elements and processes for integrating information from multiple sources. This framework must support the precise measurement of practice patterns, outcomes, and potential confounders such as severity of illness, precision of diagnosis, and socioeconomic characteristics.[18] As with the development and deployment of any information system, it is crucial that the HQIS design specifications are thoroughly analyzed through a detailed user requirements assessment (although this process is sometimes inappropriately regarded as "delaying the real work" of building the system).[26,27]
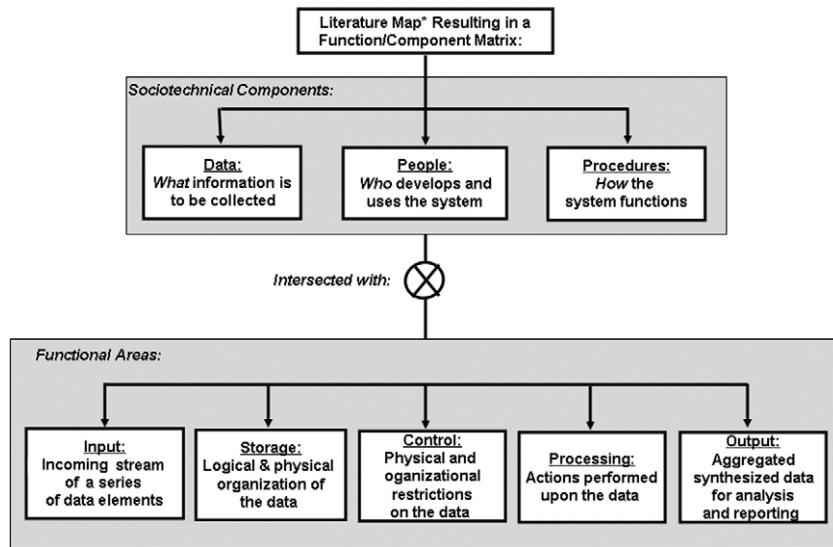
Figure 1. Literature map of the features of a successful healthcare information quality system (HQIS).

In the early stages of the health care "information revolution", technical issues received far more attention than human or organizational issues.[27] Tackling the human, organizational, and information aspects of system design requires a blend of many referent disciplines, including psychology, sociology, social anthropology, organizational behavior, organizational development, management and cognitive sciences.[27] However, a common barrier to successful design and implementation of a HQIS is an insufficient number of individuals with the necessary informatics expertise and experience within the healthcare community.[28]

HQIS human and organizational considerations include "user cordial" information system design,[29–31] end user empowerment,[32] user participation in system development,[28,33,34] and effective system implementation strategies.[35,36] Increased user involvement in the HQIS design and deployment increases system acceptance and usability, decreases resistance to change, and enhances user commitment.[37] While these "softer" factors related to information management system development may not appear difficult, in fact they are extremely challenging. Inadequately addressing these socio-technical issues is a more common cause of system failure than hardware or software deficiencies.[38]

Informatics Blueprint Construction
To identify, organize, and integrate socio-technical themes on HQIS design, development, and deployment, we applied a "literature mapping" process.[39] Such maps consist of visual renderings of the literature stemming from several referential disciplines, to assist in organizing and recognizing emerging themes and concepts, as shown in Figure 1. While conducting the literature review to create the map, the following key words were utilized in the search process:

- Healthcare Quality Field Terms: quality, healthcare quality, quality of care, quality indicators, guidelines, practice guidelines, treatment guidelines, patterns of care, outcomes, outcomes data, outcomes research, clinical research
- Informatics Field Terms: Informatics, information usage, information systems, informatics, databases, data collection, data collection systems, web, internet, internet-based systems
- Component Terms: Human, human factor, organization, organizational factor, ergonomics, socio-technical, socio-technical factor, socio-technical aspect, socio-technical characteristic

The field terms were crossed with all component terms, and the literature search was conducted using Medline, Ovid, and Science Citation Index. While we primarily focused on medical literature, as this discipline is most relevant to healthcare quality, informatics/computer science journals and texts were considered as well.

Based on the results of this mapping process we grouped common themes regarding the creation of an effective HQIS, and then applied these themes to a function/component matrix adapted from Dewitz.[40] In this paper, the purely technological components of software and hardware were not addressed; for the NCCN HQIS these have been described elsewhere).[9–10]

The first axis of the matrix describes the following nontechnical components:

- **Data** perspective, describing *what* information is of interest.
- **People** perspective, describing *who* develops and uses the system.
- **Procedural** perspective, describing *how* the HQIS functions.

The HQIS data must be carefully defined through the data dictionary development process. The people involved in the HQIS will include system users, designers, implementers, and decision-makers. In general, users and decision-makers possess high domain expertise but low technical capabilities, implementers have low domain expertise and high technical skills, and designers tend to possess a mix of domain-specific business knowledge and technical skills.[40] The procedural component of the HQIS governs the system's use, and includes formal Standard Operating Procedures (SOPs) as well as informal organizational culture.

The second axis of the matrix includes the following functional aspects:

- **Input** addresses the incoming stream of a series of data elements;
- **Storage** describes the logical and physical organization of the data;
- **Control** represents the physical and organizational restrictions placed on the data;
- **Processing** provides summary actions performed upon the data;
- **Output** consists of the aggregated, synthesized forms of data for analysis/reporting.

The input function includes all activities needed to make the data accessible for processing (e.g., select, enter, read, scan, receive, accept). Storage functions describe activities needed to maintain system data in a persistent accessible form (e.g., read, copy, update, create, transform). Control functions include the manual or automated steps required to ensure the validity and accuracy of input/output data, and the integrity of the stored data (e.g., verify, cross-check, authorize, authenticate, grant access). Processing functions describe the ways that data can be manipulated to produce value-added information, automatically or manually (e.g., sort, calculate, compare, summarize). Output functions represent those steps that produce summaries and reports of the data from the system (e.g., generate, produce, distribute, transmit, print).

Results: An Informatics Blueprint for Constructing and Deploying a Robust HQIS

When each of the three socio-technical components was evaluated within each of the five functional areas, a 15 cell matrix was created. The completed matrix formed the foundation of our HQIS "informatics blueprint" shown in Table 1, and described in the following sections.

Input: Data Perspective

To be able to merge and query data for performance and outcomes measurement, the first essential step is to thoughtfully define the types and depth of data required.[25,41,42] Not only clinical practice pattern and treatment data are needed, but also information regarding the patients' perceptions of their health, functional status, qualify of life, and satisfaction with medical care, typically obtained via patient surveys.[43] An effective HQIS should encompass the continuum of patient care, including diagnosis, primary treatment, adjuvant treatment, and long-term follow-up.

In specifying data elements to measure patterns and quality of care, Brook et al described three "data dimensions": structure, process, and outcomes.[4] **Structural data** consist of characteristics of the caregivers and institutions, such as the specialty care area, practice volume, institutional ownership, and geographic location.[44] Ensuring that all patients receive care considered to be high quality based on scientific data and expert judgment requires an assessment of process criteria.[4] **Process data** describe the encounters that take place between a physician or other healthcare professionals and the patient, such as the ordering of laboratory tests or the administration of chemotherapy. **Outcomes data** describe the subsequent health status of the patient following treatment, including administrative data (e.g., adherence to regulatory policies, treatment costs, and efficiency), clinical data (e.g., morbidity, mortality), and psychosocial outcomes of care (e.g., quality of life, spiritual well-being).[4,13,45]

Before comparisons and interpretations of optimal practice decisions are made, it is critical to capture and adjust for confounding factors that may influence the practice patterns-outcomes relationship, such as patient demographics, disease severity, comorbidity, and complexity of treatments received.[18,46] Therefore we have added a fourth data dimension representing a class of information **moderators**, data elements that help to describe for whom and under what circumstances different outcomes will be obtained. When attempting to study cause and effect, such as the link between patterns of care and health outcomes, such moderator variables are those that are antecedent or intermediate in the causal process under study.[47] For all four data dimensions, whenever possible it is preferable to record **quantitative data** that tends to be more objective, rather than more qualitative subjective measures (with the exception of patient perceptions).[41]

Input: People Perspective

In designing the HQIS, **institutional leaders** play a key role as decision-makers, outlining the goals that will guide the data collection and system functionalities. The institutional leaders need to determine the strategic intent of the HQIS, e.g., elevating care in a particular clinical area, defending the high cost of intensive but life-saving procedures such as bone marrow transplantation, fulfilling managed care contract requirements to demonstrate high quality care, etc. The nature of the stated intent(s) in building the HQIS will clearly influence the scope of data required, which in turn will influence the cost, approach, and measures of success in building the system.

Once the HQIS objectives have been clearly defined, enumeration of the required data elements follows, requiring an appropriate group of **domain experts** to specify a close-ended set of data elements that comprise the HQIS "data dictionary." The group should include individuals who best understand the content matter to be managed by the system, as well as informatics specialists well versed in ontological issues, database design, and the data access/reporting/statistical query functions to be satisfied through the HQIS. As the data dictionary construction often requires many months, it is important to budget for the requisite staff time and resources needed for this intensive taskforce work, a fact that often is under appreciated.[48]

Recently increased attention is being paid to human-computer interface design and usability evaluations of technical systems, although in many cases the practices being adopted may be somewhat inadequate.[12] Overall acceptability of a computer system is a combination of its social acceptability and its practical acceptability.[49] The system must be good enough to satisfy all the needs and requirements of the users and other potential stakeholders. Beyond practical acceptability of a new system (cost, support, reliability, compatibility with existing systems, etc.), the system usefulness to the end user is crucial for its acceptance. Nielsen[49] defines five usability attributes:

- Learnability: The user can rapidly start getting some work done with the system

Table 1 ■ Informatics Blueprint for a Successful Healthcare Quality Information System (HQIS)

FUNCTION:	PERSPECTIVE:		
	A. DATA	B. PEOPLE	C. PROCEDURAL
I. INPUT	• **Structural Data:** Characteristics of the organizations and individuals providing care • **Process Data:** Encounters taking place between healthcare givers and the patient • **Outcomes Data:** Administrative, economic, and clinical outputs of the care given • **Moderators:** Factors influencing for whom and in what circumstances outcomes differ • **Quantitative Data:** Objective information, preferred over more subjective qualitative measures	• **Institutional Leaders:** Specify the goals and objectives in creating the HQIS • **Domain Experts:** Specify and define the required data dictionary elements to meet the goals of the HQIS • **System Analyst:** Evaluates user requirements, analyzes workflow, and documents sources of information • **Application Programmer:** Creates facile user data interfaces to the HQIS • **Caregivers:** Generate the information amalgamated as data within the HQIS • **Health Information Manager:** Abstracts and codifies data from clinic notes, coordinates patient surveys, and follow-up	• **Core Dataset Creation:** Single group of data elements that meets internal management and external reporting needs • **Data Dictionary Development:** Include clear precise definition of only those elements essential to meet HQIS goals, avoiding data dictionary 'explosion' • **External Standards:** Consider pre-existing definitions and accepted ontologies whenever available • **Data Amalgamation:** Physical collection of data elements, through a manual "codification strategy, or electronically • **Systems Analysis:** Establish the data locations and flow, documented via the Unified Modeling Language
II. STORAGE	• **Patient-centric Data Structure:** Specific data on each patient integrated into a single comprehensive view of the patient • **Central Data Repository:** Data merged from various sources and across time-points for permanent storage	• **System Analyst:** Models and documents sources and flow of information • **Database Architect:** Creates information model for central repository database	• **Record Key Structure:** Allows linkage between data elements and records across multiple platforms and over time • **Retrievable Storage Format:** Data stored separately from data entry and analysis software programs, available to many users
III. CONTROL	• **Technical Metadata:** Source system, field name type, format, transformation rules, data integrity/ consistency rules, missing value codes • **Business Metadata:** Definitions, directives, synonyms, classification codes, value codes, usage within case reports forms and output reports • **Accuracy Metrics:** Cross-field logic checks, longitudinal data integrity assessments, source documentation audits, visual data review • **Completeness Metrics:** Actual vs. expected accrual, proportion enrolled, actual vs. expected forms and visits, lag time to data entry • **Electronic Audit Log:** Records to document the nature and reason for each change to the data, allowing reconstruction of the data history	• **Metadata Analyst:** Coordinates and documents data element definition process and related metadata • **Data Quality Manager:** Provides review, auditing, and training functions • **Application Programmer:** Automates logic checks, QA reports, and "suspicion checks" to be run by users routinely • **Biostatistician:** Programs cross-field and cross-record QA reports for rectification by the Data Quality Manager • **Institutional Leaders:** Determines appropriate conditions and procedures for data access and sharing • **Privacy Officer:** Carries out procedures according to the security policy • **Database Administrator:** Automates security and confidentially procedures within the HQIS	• **Metadata Synchronization:** Ensure continual alignment between data fields and the correct definitions and directives • **Data Completeness Alert:** HQIS is "self-aware" of the nature and timing of expected data elements to be completed • **Point-of-Care Data Capture:** Data at the time and place of data creation • **Automated Error Checking:** User interface traps errors via logic and suspicion checks • **Manual Data Verification:** Review of electronic data against source documents • **Change Control:** For all edits, record nature, reason, date and time, & person • **Access Control:** Protect data through authorization form, ID/passwords, digital authentication • **Confidentiality Procedures:** Restrict information to those with reason to have access, at the necessary level of data • **Data Encryption:** Encode data so that it can only be de-coded by an appropriate "key"

Table 1 ■ Informatics Blueprint for a Successful Healthcare Quality Information System (HQIS)

	PERSPECTIVE:		
FUNCTION:	A. DATA	B. PEOPLE	C. PROCEDURAL
IV. PROCESSING	• **Counting:** Generation of data on the impact of clinical guidelines on practice and quality of care • **Descriptive Statistics:** Averages and frequencies to assess on performance variances • **Normative Benchmark:** Compares distance between healthcare system's performance and a selected group mean • **Criterion Benchmark:** Compares health-care systems performance to top performer on desired measures • **Time Series:** Graphical data trends to compare the performance of a single system/subsystem to an historical pattern	• **Institutional Leaders:** Define analyses to identify trends or deficiencies • **Guideline Committee:** Expert panel to derive practical guidelines against which data are benchmarked • **Clinical Data Specialist:** Provides training in system usage, and assists in preparation of analytic data files • **Decision Support Analyst:** Analyzes HQIS data to support management decision-making and outcomes research • **Biostatistician:** Provides scientifically valid design, analysis of results, data mining for new associations, and interpretation of findings	• **End User Training:** Provide all system users and stakeholders with adequate orientation to the HQIS • **Data Query Methods:** Retrieval and exchange of data for administrative, clinical care, and outcomes research purposes, including data marts and ad hoc data queries by users
V. OUTPUT	• **Dashboard Report:** Aggregated data reflecting the input, throughput, and output variables to diagnose problem areas and derive interventions • **Patient Summary Report:** Listing of patterns of care for individual patients, to allow practitioner to evaluate the care given to each patient on a case by case basis	• **Institutional Leaders:** Design reports of data from centralized repository to inform administrators and physicians of patterns of care and trends in patient outcomes • **Caregivers:** Utilize benchmarking and patient reports to compare practice patterns to guidelines and other physicians, and to influence behavior when warranted • **Administrators:** Utilize feedback from benchmarking reports to analyze how healthcare system is performing and mandate any required changes • **Guideline Committee:** Utilize HQIS reports to assess the currency and validity of the practice guidelines	• **Report Construction:** Numeric and graphical data displays generated at specified times, with formats varied by user type and background • **Feedback Loops:** Two-way communication to practitioners and guideline committee regarding results of HQIS analysis and reporting, to improve guidelines and healthcare quality

- Efficiency: Once the user has learned the system, a high level of productivity is possible
- Memorability: The casual user can return to the system after some period of non-use, without having to learn everything all over again
- Errors: Users make few mistakes during system use, can easily recover from any errors they do make, and catastrophic errors do not occur
- Satisfaction: Users are subjectively satisfied and find it pleasant to use the system

An essential human resource to ensure adequate system design, acceptability, and usability is the **System(s) Analyst**, an expert who evaluates the user requirements, analyzes the workflow, and documents sources of information. This helps to ensure usability by making the information and workflow as efficient as possible, leveraging any existing electronic data, and avoiding redundant data. Another human resource necessary for effective system acceptability and usability in inputting information is the **Application Programmer,** to create facile data interfaces to the HQIS that can be navigated by the user with ease. The interfaces may consist of data entry screens, or an export-import routine to capture pre-existing data from an electronic source to populate the HQIS data repository.

An important constituency involved in the success or failure of the HQIS consists of the **caregivers** themselves, as the ultimate quality of HQIS data begins with the diligence of the individuals who generate it.[41,50] As participants and facilitators in the healthcare quality assessment process, caregivers should be apprised of the need to obtain complete accurate treatment and outcomes data from the information they generate. While direct entry of coded data into the HQIS by caregivers is ideal, this approach is very difficult to achieve for a number of reasons, such as the increased time required during a patient visit.

A practical alternative to point-of-care coded data capture by the caregivers themselves continues to be the employment of highly trained **Health Information Managers** to provide accurate data collection and coding.[13] The American Health Information Management Association (AHIMA) has defined the role of the Health Information Manager as an individual who abstracts data from medical records, and other available sources, for entry into the HQIS. If standardized information can be provided by the caregiver in the medical record, data abstraction based on these records will be much more accurate.[51] In cases where the data are best obtained via direct patient survey, the Health Information Manager ensures timely dissemination and return of the surveys for encoding in the HQIS, and contacts patients when they are due for regularly scheduled follow-up.[41]

AHIMA has set forth a model curriculum that identifies the path to developing such knowledge among staff members.[13] The association also has outlined the ideal background for the Health Information Manager in outcomes research, which includes knowledge of qualitative and quantitative research methods, data structure, database management, as well as some familiarity with statistical data analysis and information systems development.[13]

Input: Procedural Perspective

In selecting the data to be collected, the goal should be the creation of a single, **core dataset** that meets the needs of both internal management and external reporting.[52] **Data dictionary development** through data element selection is a critical task. The total number of selected variables should be considered to avoid the phenomenon of "data dictionary explosion," as too large a database can prove cumbersome to manage and costly to maintain. In addition, an inverse relationship between quantity and quality of data has been observed, such that, as the effort required to complete a data form increases, the quality of the information frequently decreases. Thus it is crucial to discriminate between essential and nonessential data elements, and to avoid redundancy.[41] However balance is required, as an overly limited data dictionary may compromise the ultimate power of the analyses.

In developing the data dictionary, a key goal is to optimize the capacity for data sharing across systems and studies. Obtaining this goal requires the clear specification of a precise consistent definition for each data element identified in the data dictionary, aligning with pre-existing **external standards** within accepted national/international ontologies whenever they exist.[51] To enable data sharing across multiple hardware and software platforms, standardized common definitions ideally should be utilized across and within centers.[25,53] A number of national/international efforts are underway to develop such standards, such as the National Cancer Institute's Cancer Biomedical Informatics Grid (caBIG).[54]

While this standardized approach represents the ideal, currently a lack of widely accepted common vocabularies across disciplines and settings hampers this process.[48] There is a high degree of variability inherent in biomedical vocabularies, and controlled vocabulary systems are in a continual state of development, expansion and refinement.[55,56] Yet if a HQIS utilizes heterogeneous or proprietary vocabularies, this will lead to incompatible semantics and the inability to integrate information across systems.[57]

Emerging semantic standards for medical data include the Systemized Nomenclature of Medicine (SNOMED) created by the College of American Pathologists (CAP), one of the most robust reference terminologies for medicine.[58–61] For laboratory results data, the Logical Observation Identifiers Names and Codes Ontology (LOINC) is recognized as the most comprehensive standardized vocabulary.[62] While no ideal standard vocabulary currently exists for coded drug data, RxNorm is a very promising drug vocabulary in development by the National Library of Medicine, in consultation with the Food and Drug Administration, the Department of Veterans Affairs, and the Health Level-7 standards development organization.[63] SNOMED, LOINC, and RxNorm are recommended standards for the future national Electronic Health Record System (EHR-S); therefore selecting these standards for a HQIS provides the best opportunity to re-use clinical care data collected within the EHR-S for healthcare quality and outcomes research.

The term **data amalgamation** has been used to describe the physical collection and retrieval of data elements set forth by the data dictionary, leading to the storage of these data in a

data repository.[64] These data may arise from a number of sources, including: medical records; department quality assessment and improvement processes; case management, utilization management, and risk management systems; patient satisfaction surveys; and cost accounting systems.[13] Data amalgamation can be achieved through either manual or electronic data retrieval. In the manual process, the required data are realized through a "codification strategy."[64] This process entails the management of knowledge through a "people-to-documents approach", in which information provided by someone with the requisite knowledge (e.g., a physician who dictates a treatment note in the chart) is then coded by another individual, making the data available for reuse in various systems, independent of the originator of the information.

In healthcare quality assessment, codification involves examination of caregiver's clinical notes by the Health Information Manager, and extraction of this information into coded data fields. To facilitate this process, various data interfaces should be considered, including web-based data entry screens and scannable forms, clinical data capture methods that have been shown to produce gains in efficiency and accuracy.[65] A mix of system interfaces may be optimal to capture treatment and outcomes data generated from various care settings. For example, scannable forms might be utilized for data collected in a busy clinic area with limited Internet access; a web screen could be provided for entry of data acquired by the Health Information Manager through review of source documents in the office setting; and paper survey forms with a stamped self-addressed envelop might be mailed to patients to obtain follow-up information.[51]

Automating data acquisition to the greatest extent possible can reduce errors, labor, and expense when compared with abstraction onto paper records. Because healthcare quality assessment focuses on routine care practices in patients representing the entire patient population, data on these patients typically will not have been acquired in a research system, but rather reside in a computerized or paper-based medical record. Ideally, the HQIS should communicate with external databases to receive electronic data appropriate for measuring quality of care and outcomes whenever feasible.[66] However, if electronic data retrieval is to be a viable option to amalgamate the needed information, such preexisting data must have been collected under definitions and rules that are compatible with the HQIS data dictionary.

As it is unlikely that any one information system will contain all of the requisite information, a full **systems analysis** is needed to establish the data acquisition flow, determining whether and where the required data currently exist in an appropriate form.[41] Identifying the many varied department and organization-specific sources of useful information for outcomes research is a daunting task.[13] In conducting this systems analysis, the information resources and requirements should be documented utilizing the Unified Modeling Language (UML), a standard and widely accepted modeling approach that assists in specifying the internal operations and data structures to facilitate subsequent application and database development.[67]

While electronic billing or pharmacy claims databases may be an attractive means to facilitate healthcare quality assessment, such systems often do not communicate with one another readily, so that accessing the information on a large scale and in a timely cost-effective manner can be quite challenging.[13,41,68] For the primary purposes of capturing charges, generating bills, and collecting payment, patient information within the EHR-S often is organized around the encounter, rather than the patient, making it more difficult to tap into this resource for the purpose of the HQIS.

Storage: Data Perspective

To support the HQIS goals of measuring and assessing quality of care and outcomes, the manner in which data are stored within the repository is fundamental to successful data usability. To support systematic improvements in healthcare and outcomes, a **patient-centric data structure** is required, in which records within the database are organized around the patient, and can be linked together via a patient-specific ID such as medical record number.[69]

Some systems lend themselves to "federated" databases, in which the data are not pooled in one physical location, but rather reside at each local institution generating the data, while being made accessible and searchable by other institutions and users.[70] However the HQIS is more amenable to establishing a **central data repository** as the persistent database layer. The data required for the HQIS accumulate over time in a longitudinal fashion within the patient care records, and only the subset of data specified within the dictionary needs to be retained in the centralized repository. Furthermore patients may transfer and receive their care from a number of different organizations over time. When data reside in multiple non-centralized repositories that are changing over time (as in the federated approach), continual linking and re-extraction of the relevant subset of data becomes more difficult to manage than storage within a single central repository.

Storage: People Perspective

A key human resource to facilitate appropriate storage of the HQIS data is the **Systems Analyst**. This individual is responsible for gathering the user requirements, information flow, and workflow that will guide the construction of the HQIS database tables and relationships for ultimate data storage. The **Database Architect** is the technical expert who then uses these specifications to create a robust information model underlying the central data repository.

Storage: Procedural Perspective

An underlying premise behind many important healthcare delivery initiatives is that patient data, collected over time and stored in one or more information systems, can be accurately integrated into a single, comprehensive view of the person.[25] To do so successfully, it is critical to establish the correct **record key structure** for linkage between data elements and records. This requires that the key identifiers, and any secondary identifiers, be appropriately established to provide the capacity to access, move and integrate among databases for a given patient's data elements and records.[25] For example, a HQIS repository may utilize medical record number of the patient as the primary key, to allow linkage of multiple data records on the same patient, and visit number

as a secondary key, to allow sequencing of information on a given patient.

Data storage should enable mining of the repository for administrative, quality of care, and outcomes research purposes. Careful attention should be paid to creating appropriate relationships among data to facilitate data retrieval and manipulation. Patient-specific data should be entered in a **retrievable storage format** within the central repository, rather than embedded within the software applications that facilitate data entry and analysis, such that incoming data on patients is stored distinctly from the application programming itself.[35–37] The data must be stored in the central data repository in such a way that the data are entered once, yet are available for use by many individuals in a number of different ways.[71,72]

Control: Data Perspective

To successfully conduct healthcare quality assessment and outcomes research, the validity, accuracy, and completeness of the data are crucial factors. To document the meaning and structure of the data collected within the HQIS, in addition to the raw data, it is critical to collect and store the appropriate "metadata" (defined as "data about the data"), to describe and explain the information being collected, managed, and stored through the HQIS. The metadata make it possible to retrieve, utilize, and manage data as an information resource, and to conduct accurate useful analysis and reporting of the HQIS data.[73]

Metadata can be categorized as either technical or business metadata.[74] **Technical metadata** represent information needed from the database administrator viewpoint, such as: source system; field name, type, and format; transformation rules for derivation of values; rules for data integrity and consistency; and missing value codes. **Business metadata** describe the data from the user's perspective, such as: the data meaning (definitions); instructions for data collection (directives); synonyms and classification codes to facilitate searching for data elements; allowable value codes; and cataloguing the specific case report forms and output reports that use each data term. While the technical metadata arise from the construction of the physical database and are stored within the application, the business metadata are extracted from the domain experts, and often are handled in a separate, linked database or even spreadsheet. In future, systems that link and store technical and business metadata together in association with the data repository would be ideal.

The HQIS must accommodate and provide for methods to ensure data quality, and to detect and reject data records of low quality. At least two types of metrics are required to measure the quality of the acquired information: accuracy metrics and data completeness metrics. **Accuracy metrics** help to ensure that the information collected is correct, and include measures such as the results of cross-field logic checks, longitudinal data integrity assessments, audit results comparing submitted data against source documentation, and the outcome of visual data review. Data **completeness metrics** help to avoid information gaps or missing data. Such metrics provide indices of the completeness of patient coverage within the repository (e.g., actual versus expected patient accrual), gaps within an individual patient's infor-

mation (e.g., expected versus actual data forms and visits), and lags in data acquisition within the HQIS (e.g., difference between data expectation and data entry dates).

Another critical set of records within the HQIS is the **electronic audit log**, documenting the identities of all users of the data, and exactly which data were utilized or accessed. The audit log documents the nature and reason for each change to the data, which allows for reconstruction of the entire history of the data. An electronic audit log also should facilitate the ability to produce exception reports and translation logs of any rollback procedures, and provide maintenance reports to identify such issues as missing key values.

Control: People Perspective

Technical metadata usually are generated through programming of the database management system itself through an automatic process.[73] However the business metadata must be carefully documented and maintained through human organizational processes. In addition to the domain experts who assist in data dictionary development, a designated **'Metadata Analyst'** should be appointed. The Metadata Analyst is an information professional trained to coordinate the data element definition process, and to maintain carefully documented business metadata throughout the life of the HQIS.[73] It is generally best if the Metadata Analyst work closely with those who originate the data (e.g., the caregivers) to obtain the needed data definitions, as the data originators have significant understanding of the rationale for the dataset and its potential uses.[73] While this process may represent the best practice, it may be more efficient to have the Metadata Analyst create the definitions for subsequent review by the creators of the data, who often do not have the time or skills to create the business metadata themselves.[73]

With respect to data quality controls, individuals with the requisite content knowledge, logical orientation, and "people skills" are required to most effectively conduct continual data quality assessment. One of several emerging roles for Health Information Managers as defined in AHIMA's "Vision 2006" report includes that of **Data Quality Manager**.[13] The Data Quality Manager provides regular review and auditing training functions to continually assess and improve data quality. In addition, the Data Quality Manager specifies the quality assurance algorithms and reports that may identify anomalies or errors not captured during logic checking of the data as they were entered. Individuals who themselves have experience in data collection, and who possess a strong grasp of the domain content, are best equipped to fill this role.

Once the quality assurance checks and reports have been fully established, they can then be automated for optimal efficiency by an **Application Programmer**. This allows quality assurance checks to be invoked at the point of data entry, e.g., by providing an error message if a datapoint entered is beyond the allowable range for a field. Alternatively logic check facilities can be built into the HQIS to be run by the user at frequent intervals, e.g., a menu item that screens all data entered on a given patient for logical sequencing of event dates. Computer-generated cross-field and cross-data record logic checks and data quality "suspicion" reports also can be generated by the **Biostatistician** who will be analyz-

ing the data, as data screening is carried out prior to analysis. The reports would be reviewed by the Data Quality Manager working with the Health Information Manager, to either confirm or correct the suspect results. This process not only produces the highest quality data, but also confers the benefit of familiarizing the Biostatistician with the data structure and content in preparation for the statistical analysis.

In granting access to the HQIS repository, **institutional leaders** (including domain experts as well as administrators) should establish the policies and procedures dictating the appropriate levels of security and data access among the members of the research team and system users. This group should determine and publish an appropriate data sharing policy that clearly defines the circumstances under which data can be shared across the participating institutions. This policy should be established early on, before amalgation and knowledge browsing of data occur.

In a world regulated by the Health Insurance Portability and Accountability Act (HIPAA), a **Privacy Officer** may be appointed to enforce the security policies and grant system and data access to appropriate individuals following the established procedures.[51] The **Database Administrator** is the individual responsible for incorporating these security and confidentiality rules into the system, automating the processes as much as possible, and providing an audit trail to track that security and confidentiality are being appropriately maintained.

Control: Procedural Perspective

Control processes should be deployed to ensure that unauthorized access is not permitted; all data are captured, stored, transmitted, and displayed without error; and audit trails of all transactions exists. A key control process charged to the Metadata Analyst is **metadata synchronization** to ensure that there is continual alignment between the data fields and the correct definitions and directives, documented in a metadata repository.[73] It is essential to thoroughly document how the data collection and storage strategy evolves over time, as is inevitable, to minimize problems of converting corrupted and missing fields, and to inform the analyses conducted by the bio-statistician.[41]

A growing number of free and commercial metadata tools are available, such as templates that allow a user to enter metadata values into pre-set fields and to generate a formatted set of data element attributes and corresponding values.[73] Mark-up tools structure metadata attributes and values into the specified schema language, generating Extensible Mark-up Language (XML) Document Type Definitions (DTDs). Extraction tools automatically create metadata via analysis of a textual digital resource; however as the quality of the extracted information will vary based on the tool's algorithms and the source text content, this tool best serves as an aid to creating metadata, still requiring manual review and editing by the Metadata Analyst. There also are conversion tools that translate one metadata format to another, again requiring manual review and editing depending on the degree of similarity between the metadata elements in the source and target formats.[73]

For valid healthcare quality assessment and outcomes research, the task of perfecting a database is an ongoing one that is never fully complete. An independent, comprehen-

sive data validation process should be put into place for the HQIS, such that the same high standards expected of any prospective epidemiologic study are met.[25,75] An efficient HQIS should be "self-aware" of the nature and timing of the data elements expected throughout the data amalgamation process, and should prompt the user when it is time to collect data.[76] This form of **data completeness alert** will help ensure that all required records and data elements are captured. To obtain the most accurate complete information, **"point-of-care" data capture**, in which acquisition of patient data occurs as close as possible to the time and place of its generation, is highly desirable (although difficult to achieve with the current state of the EHR-S).[51,72]

Regardless of the mode or source of data acquisition, the HQIS human-computer interface should provide built-in mechanisms for **automated error checks**, to "trap" errors at the point of data entry based on algorithms provided by the domain experts. "Logic checks" are invoked when the right and wrong answers can be specified for a given data element. "Suspicion checks" are used to point to data that may be erroneous based on outlying values or improbable combinations of data, for further visual review and investigation. With scannable forms, automated coding verification routines can be incorporated into the data scanning process, to check for and disallow invalid data at the point of entry.[41] Similarly in a web-based data entry interface, logic and range checks can be invoked within the data entry screens to avoid entry of erroneous data.

In spite of automated error checking processes, some **manual data verification** will still be required. Scannable forms attempt to recognize handwritten text, yet they rely on the Data Quality Manager to confirm that correct data are transferred to the computer.[41] As automated logic checks cannot trap those errors that appear to be within reason based on error-checking algorithms, a source document review needs to be conducted to confirm the data accuracy.[41] In this process, a randomly selected subset of the electronic data records stored in the repository, and/or the most critical data elements, are compared back to the original source documents from which the information was generated, through a manual review.

To ensure that only appropriate changes are made by authorized individuals, a **change control** process should be implemented. An electronic audit log should record the nature of the change, the reason for the correction, the date and time of the change, and the individual making the change. Such an audit trail will allow for identification of any recurring problems by reconstructing the history of the editing process at any time.

As electronic data storage makes information readily available but also vulnerable to unauthorized use, it is critical to maintain strict **access control** data security enforcement.[25,41] Particularly in this era of HIPAA regulations, for any biomedical database, policies and procedures carefully outlining the conditions and terms under which user access to data will be provided need to be in place, including an authorization form that is compliant with HIPAA.[16,38,41] ID and password protections should be invoked within the HQIS, along with digital authentication of the individual

requesting access, to ensure that the user is who he or she claims to be.

Role-based security processes can achieve protection of the information through restriction of access to the HQIS itself, and/or restricting access to particular types of information within the system. **Confidentiality procedures** restrict access of information to only those with appropriate reason to have such access, and to only the necessary level of patient identification. Data sharing policies should be in place to define appropriate data sharing across institutions under HIPPA compliant conditions. Before any pooled data are released, individual patient identifiers should be stripped, or **data encryption** utilized to encode the information in a manner that can only be de-coded by an individual holding the appropriate "key" to the encryption process.[66]

Processing: Data Perspective

However sophisticated the data capture, and reliable/valid the measures, the HQIS must be designed to allow processing of the data for maximum usefulness. For healthcare quality research, patient-specific data must be made available in a suitable format to facilitate treatment decision-making. Benson has used the term **"counting"** to specifically refer to the use of informatics to generate and analyze data about the impact of clinical guidelines on practice and quality of care.[77]

Data usage also can be termed "knowledge browsing", with the most basic form being the construction of **descriptive statistics** (averages, frequencies) that can serve as a basis to provide some information regarding variances in healthcare performance.[78] A second form of healthcare quality knowledge browsing consists of benchmarking, i.e. employing a standard or target to judge best performance in a group.[78] Such cross-sectional comparisons can be used to assess the performance of one healthcare system or organization against others to determine whether they are similar.

A **normative benchmark** reflects a healthcare system's performance in contrast to that of a selected group. This benchmark can be constructed as the distance of a healthcare system's performance from the group mean, and then determining whether this distance falls within some acceptable level of variance, e.g., within 15%. A **criterion benchmark** reflects a preset, desired performance level (e.g., productivity, cost targets). When a system's performance is compared against the top performer in a group or a specified performance target, this is known as a "best-in-class" benchmark or comparison.[77–80] while the criterion benchmark typically remains somewhat constant over time, the normative benchmark is more likely to vary across measurement time intervals. **Time series** comparisons, or data trending, can be used to compare the current performance of a single system/subsystem against its historical pattern or mean performance. Such trending data usually are graphed over time to provide information about the general stability of system activity and performance, or insight into seasonal changes over a given year.

Processing: People Perspective

In assessing healthcare quality, **institutional leaders** must take a key role by defining the usage of the data in forms of knowledge browsing and benchmarking. Investigators should outline the concepts for guiding the outcomes re-

search and the data analysis to be performed from the centralized pool of HQIS data, assisting the institution to identify time trends or possible deficiencies in the patterns of care. Typically a **Guideline Committee** consisting of a panel of expert physicians is established to derive practice guidelines, against which the data accumulated in the HQIS will be compared.[81] These guidelines represent a statement of consensus of a group of domain experts regarding their views of currently accepted approaches to treatment, based on clinical trials evidence published in the medical literature.

It has been found that all manner of research staff could benefit by becoming better prepared to assume leadership roles in conducting outcomes research.[13] Two emerging roles defined in the AHIMA's Vision 2006 to support knowledge browsing within an outcomes research HQIS include Clinical Data Specialist and Decision Support Analyst.[82] The **Clinical Data Specialist** provides system training to the users of the HQIS and assists with preparing for proposed analyses, thus empowering the users and further enhancing the system usability. The **Decision Support Analyst** supports senior management by providing the information needed for decision-making and strategy development out of the HQIS, using a variety of analytical tools.

A natural background for an individual assisting in the interrogation of a complex data repository is that of **Biostatistician,** who should be a full collaborator in the research process, not merely someone who consults on or controls this process.[83] A medical statistician's routine professional activities can have important ethical consequences, as scientifically valid medical research requires high quality statistical design and data analysis as precursors.[84,85] As the volume of available complex biological and treatment data increases through the HQIS, the statistician will apply data mining tools to search for previously hidden associations between patterns of care and outcomes. As a collaborating scientist in the outcomes research, the statistician should ensure that the interpretation of the findings from the HQIS and the medical decision making conform to sound data integrity and statistical principles.[84]

Processing: Procedural Perspective

By improving providers' access to healthcare quality data, access to the HQIS will promote future research to maximize the value of the health care system.[86] For successful use of an HQIS, it is crucial to support new adopters of the database system with **end user training**. To do so it is necessary to allow for sufficient time and provide adequate staff to effectively train all system users and stakeholders.[48]

If properly structured, the information framework of the HQIS creates the capacity for the user to apply **data query methods** to the resulting database, and to retrieve and exchange data between authorized entities for administrative, clinical care, and outcomes research purposes.[25,86] An effective HQIS will allow data querying either directly, or via files download to standard data analysis packages.[35,37] Such knowledge browsing will generate outputs in many forms, including static reports, algorithms for data aggregation and analysis, "data marts" of pre-aggregated information, and data queries issued directly against the source data.

To encourage exploration of the data by investigators, facile user query capabilities should be built into the HQIS, including ad hoc query capabilities.[87] Ad hoc querying will include queries issued directly against the source data (or a copy with the same structure residing on a server different from the transactional system), and queries of data subsets that have undergone major restructuring and importing into specialized database management systems optimized for query, such as multidimensional database engines (data marts).[87,88]

Ouput: Data Perspective

The objective of an HQIS is to measure practice patterns and to benchmark medical practice against clinical practice guidelines to enhance clinical decision-making, with the ultimate goal of improving quality of care. To do so will require that effective reports be constructed to describe healthcare performance data. One form of a highly useful summarized reporting tool is the **dashboard report**. This report consists of aggregated data that reflect the input, throughput, and output variables of the system/subsystem. The display of data within the report should be made quickly informative of the system's performance by utilizing varied formats such as line graphics, bar charts, pie charts, etc. Review of a dashboard report should help institutional leaders diagnose problem areas within a healthcare performance system, and to derive interventions for healthcare performance improvement when needed.

A more detailed form of **patient summary report**, describing the patterns of care for individual patients, also is a highly useful form of information for the caregivers themselves. These reports allow the practitioner to evaluate the care given to each patient on a case by case basis, forming a feedback loop that may lead to performance improvement in future care.

Output: People Perspective

As data are benchmarked against the established care guidelines, it is essential for **institutional leaders** to learn to use the HQIS by critically reviewing the output reports (which may sometimes include conflicting information), to "read the signposts", recognize trends in time, and envision a plan and respond to any negative trends in an effective manner.[48] The **caregivers** for the patient population under study require feedback reports from the benchmarking process, to inform them of how their practice patterns compare to national guidelines and other physicians, and to influence their behavior when warranted. In addition, **hospital administrators** require reports to determine how well the overall healthcare system is performing, and to mandate changes if required. Depending on how they adopt to utilizing the HQIS reports, an administrator may serve as either a leader or a deterrent for an organization to successfully adopt performance improvement.[48] Finally, the **Guideline Committee** members who created and maintain the practice guidelines for the best standard of care require feedback from the HQIS reports, to allow them to assess the currency and validity of the guidelines.

Output: Procedural Perspective

Having created the HQIS, it is usual to put in place routine **report construction** of outcomes data, and to generate these reports for review at specified time intervals. In creating data displays and reports for users it is important to note
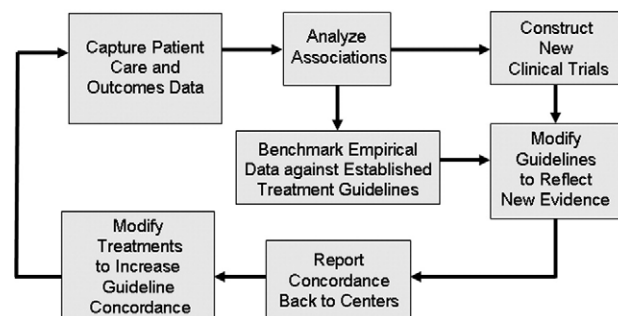

F I G U R E 2. Feedback mechanisms to optimize guideline consensus and clinical decision making.

that the data presentation may need to be varied by the type of user and their particular background, to ensure that the information is imparted efficiently. A highly useful form of data display will incorporate graphics (e.g., bubble charts depicting length of stay and cost by practitioner).[76]

One way to improve healthcare performance is to bring a level of accountability to clinical practice.[89] The processes of data manipulation and knowledge browsing can produce valuable information that may lead to new knowledge. For example, examining observed patterns of care that are highly disconcordant with prevailing guidelines may lead to the knowledge that the guidelines require modification to accommodate key factors such as extreme comorbidity; alternatively, investigating such patterns of non-concordant care may lead to the knowledge that additional physician training and decision support are required to improve guideline concordance. Such information must be fed back to the appropriate constituencies to facilitate uncovering the knowledge that can effect change and improve patient care. This requires creating two-way **feedback loops** both to the practitioners, to inform them of how their practice patterns benchmark against national guidelines and other physicians, and to the committee developing the guidelines to ensure that maximum currency and consensus is achieved.

Figure 2 depicts the information flow within an outcomes research and performance improvement program, to effect guideline modifications and changes in practice when deemed appropriate based on patterns of care analyses. New studies may have a major impact on practice recommendations. If mechanisms do not exist to rapidly incorporate this information into the pathway, the entire guideline program will lack relevance. In addition a feedback loop is needed to incorporate the results of guideline concordance analyses that may indicate that thought leaders have adopted emerging new standards of care, ahead of the medical literature publication process. A lack of consensus demonstrated by these analyses could indicate the need for further refinement of the guidelines, to take into account special circumstances or populations.[81]

Communication strategies are needed to close any gaps in the feedback loop to the healthcare institutions and providers. A challenge for the HQIS design is to create faster, condensed, organized, and more accurate feedback to support decision-making, without overloading users with information.[48] This information should include the guideline concordance results and summaries of patterns of non-

concordance, to allow performance improvement activities to take place at the caregiver level when appropriate.

Discussion

In a 1997 Nursing Informatics conference on Patient Guidelines and Clinical Practice Guidelines, the most critical issue identified with respect to healthcare improvement was that of identifying data that could be used to examine clinical practice variability, and thereby establish evidence for improving the quality of clinical practice.[90] Care guidelines, protocols, and clinical pathways have evolved in recent years, in an attempt to standardize patient care, reduce complications, decrease length of stay, and improve outcomes.[1] Yet very little information about the "processes of care" is evident, primarily due to a lack of development and deployment of an effective HQIS within healthcare institutions.

The new focus for lowering healthcare costs, while maintaining or ideally improving quality of care, demands routine measurement of practice patterns and outcomes. Over the past 50 years, various forms of computer-based information management applications have been developed and deployed in the clinical setting. However additional informatics tools are necessary so that a convergence of clinical data with evidence-based inquiry can occur.[50,90,91] Measuring healthcare quality without such automated tools is extremely time-consuming and labor-intensive; therefore it becomes crucial to support the creation of effective information systems on healthcare performance measurement.[17] Through the HQIS it will become possible to perform quality measurement in ways that will be less expensive, yet more comprehensive and reliable, than previous methods.[17]

While the development of Internet-based systems, open architectures, and data exchange standards have begun to address technical barriers to outcomes research HQIS development,[13] similar advances in the socio-technical and KM factors that underlie a successful system to measure and improve the quality of care are needed. To improve future patient care, it is extremely important to guide physicians, researchers, administrators, and informaticians in understanding the optimal socio-technical and KM processes and resources needed to create a successful HQIS.

It is this need that led to the rationale for developing this informatics blueprint for the design and deployment of a robust HQIS. The socio-technical challenges associated with capturing useful data concerning patients' perceptions of their own health status, physician practice patterns, and valid outcome measures are great. This drives the requirement not only for better information systems, but also for advice regarding the related processes and human resources to allow capture of the needed outcomes data, and to ensure that such data are properly coded. However it must be kept in mind that the ultimate success of any system lies in its effective acceptance, adoption, and utilization by the organization for which it is built. In the case of the HQIS, a number of organizational competencies and a certain organizational culture need to be in place to ensure the success of the system. These include, but are in no means limited to: an appropriate high level institutional official to champion and drive the initiative and be held accountable for its failure if this is the result; highly skilled Systems Analysts and other

experts to investigate and resolve the complex underlying information sources and workflow to make the data collection and reporting processes as efficient as possible; and support and buy-in from the caregivers whose performance is being captured and in some senses rated.

No doubt outcomes research systems will continue to evolve over time, and fundamental shifts are occurring that will facilitate the collection and assessment of healthcare information.[13] Paper-based human collection and keyboard input are being supplanted at a rapidly increasing pace, spurred by the reduction of costs of electronic processing and image capture technologies, the maturing of the Internet, the introduction of handheld devices, direct electronic capture data, and voice input of information.[69] The human resources and organizational processes described in our informatics blueprint will need to be adopted and evolve in parallel with these technological advances.

While it is often hoped and predicted that the adoption of new technologies will result in a reduction in human resource costs, to date the experience is that staff reduction levels are not actually achieved.[12] Methods to assess organizational and user acceptance will be necessary if emerging forms of technology are to be effectively deployed. Performance measures such as output (productivity, goal achievement) and error or quality assessments are critical; however, these techniques are not useful in assessing usability or acceptability of new systems.[92] Successful evaluative techniques will be those that can record the implicit cost-benefit assessment made by users every time they choose to use (or not use) a particular facility of the system.

Outcomes data are quite rich and can serve many purposes: clinical and administrative operations management, adherence to regulatory policies, marketing and research.[76,93] Yet much of the data needed for outcomes studies either have not been collected, or cannot be analyzed in their present format, as patient information across encounters and facilities is difficult to retrieve.[13] The development and refinement of outcomes research programs will be aided immensely by the creation of comprehensive outcomes datasets through the HQIS. These systems must be able to manage enormous amounts of data, and facilitate the analysis of multi-faceted information, including administrative, financial, and clinical data.[13]

Because clinical outcome studies incorporate the imperfections within human nature, including errors or gaps in knowledge abstraction and coding, it is important to acknowledge the inevitable shortcomings in the data acquired for assessing quality of care through the HQIS.[41] Brook et al emphasized that it will never be possible to produce an error-free measure of the quality of care.[4] However, poor measures of quality can unfairly harm institutions and physicians, such that every effort should be made to use state-of-the-art measures, even if additional expenditure is required. While comparative data should motivate a physician to examine how to improve the care provided, practitioners will not use such comparisons unless there is an appropriate accounting for differences in patient populations.[76] Such risk adjustments are required to make valid comparisons, and with the proper data dictionary, this complex task can be facilitated by an information system.[76] The roles described in the informatics blueprint for an

outcomes research HQIS will evolve over time as well. AHIMA's Vision 2006 defines a number of emerging roles for health information managers that directly relate to the outcomes movement.[13]

Feedback loops are included in the informatics blueprint as a critical prerequisite for an effective clinical decision-making process based on an HQIS. The absence of clinicians and patients in the feedback loops of attempts to measure care and its effects is one of the greatest weaknesses that detracts from the quality effort.[90]

Internet-based data collection efforts are becoming more common, and have been used for benchmarking patient care through outcomes research.[94–97] Deployed in 1997, the NCCN Outcomes Research database represents one of the earliest such systems. This HQIS was targeted specifically at overcoming the deficit in linking treatment patterns with patient outcomes, by providing the capability to collect and report patterns of care and outcomes data in the oncology setting, via an Internet-based data system.[98] Combining the familiarity of the Web environment with its cross-platform compatibility and real-time data access and submission capabilities, the NCCN system has been readily adopted by 15 participating centers nationwide to date. Secure transmission of data and guideline concordance reporting has been achieved for the past 10 years, and pattern of care reports have been used in a variety of ways by participating NCCN centers. In user surveys, the majority rated the system as "good" to "excellent" with respect to layout and flow of the Web site, ease of downloading documents, ease of use as a database entry system, usefulness of on-line logic checks, and ease of use as a database reporting system.[98] Now that our informatics blueprint for an HQIS has been created based on an extensive literature review, a more extensive assessment of the NCCN HQIS against the full blueprint recommendations is underway.[76] This "case study" applying the healthcare quality information system blueprint to the NCCN system will help to assess the utility of the blueprint in the construction of future systems.

Summary

We believe that the informatics blueprint presented here, drawn from the experiences of numerous outcomes research initiatives and domain experts, could serve as an extremely useful guidepost to developing an outcomes research HQIS. This focus on the socio-technical and KM components of building such a system is one that often is overlooked or under appreciated, in preference to emphasis on the more enticing technological aspects. While some components of the blueprint are specific to outcomes research, much of the advice it contains may well be applicable to the planning, development, and evaluation of data systems developed for other areas within the clinical research arena, and perhaps beyond.[79,80]

References ■

1. Hammond JJ. Protocols and guidelines in critical care: development and implementation. Curr Opin Crit Care Dec 2001;7(6): 464–8.
2. Craft PS ZY, Brogan J, Tait N, Buckingham JM. Implementing Clinical Practice Guidelines: A Community-Based Audit of Breast Cancer Treatment. Med J Aust 2000;172:213–6.
3. Richards M SR, Kent D. Inequalities in breast cancer care and outcomes. Br J Cancer 2000;76:634–8.
4. Brook RH, McGlynn EA, Cleary PD. Quality of health care. Part 2: measuring quality of care. NE J Med Sep 26 1996;335(13):966–70.
5. Simmons BP SM, Evans RW, Amadio PC, Cats-Baril W. Outcomes assessment in the information age: Available instruments, data collection, and utilization of data. Instr Course Lect 1999;48:667–85.
6. Ritter MA AM. Overview: Maintaining outcomes for total hip arthroplasty: The past, present, and future. Clin Orthop Relat Res 1997;344:81–7.
7. Epstein AM. Rolling down the runway: the challenges ahead for quality report cards. JAMA Jun 3 1998;279(21):1691–6.
8. Chassin MR, Galvin RW. The urgent need to improve health care quality. Institute of Medicine National Roundtable on Health Care Quality.[see comment]. JAMA. Sep 16 1998;280(11): 1000–05.
9. Niland JC. NCCN Internet-based data system for the conduct of outcomes research. Oncology (Huntington). Nov 1998;12(11A): 142–6.
10. Niland J. NCCN Outcomes Research Database: The First Two Years of Data Collection and Analysis. Oncology. 1999;13(11A): 575–8.
11. Weeks J, Niland J. NCCN Oncology Outcomes Database: An Update. Manag Care Cancer May/June 1999:25–32.
12. Eason K. Changing Perspectives on the Organizational Consequences of Information Technology. Behav Inf Technol 2001;20: 323–8.
13. White AW, Wager KA. The Outcomes Movement and the Role of Health Information Managers. Top Health Inf Manage May 1998;18(4):1–12.
14. Smith TJ, Hillner BE. Ensuring quality cancer care by the use of clinical practice guidelines and critical pathways. J Clin Oncol Jun 1 2001;19(11):2886–97.
15. Hillner BE, Smith TJ, Desch CE. Hospital and physician volume or specialization and outcomes in cancer treatment: importance in quality of cancer care. J Clin Oncol Jun 2000;18(11):2327–40.
16. Clancy CM, Lawrence W. Is outcomes research on cancer ready for prime time? Med Care Jun 2002;40(6 Suppl):III92–100.
17. Bates DW, Pappius E, Kuperman GJ, et al. Using information systems to measure and improve quality. Int J Med Inform Feb-Mar 1999;53(2–3):115–24.
18. Pilote L TI. Outcomes research in the development of practice guidelines. BMC Health Serv Res 2002;2(1):7.
19. Humpries KH RJ, Carere RG, Buller CE, Kiely FM, Spinelli JJ. Co-morbidity data in outcomes research: are clinical data derived from administrative databases a reliable alternative to chart review? J Clin Epidem 2000;53(4):343–9.
20. van Walraven C WB, Ugnat AM, Naylor CD. False-positive coding for acute myocardial infarction on hospital discharge records: chart audit results from a tertiary centre. Can J Cardiol 1990;6(9):383–6.
21. Iezzoni LI. Assessing Quality using Administrative Data. Ann Intern Med 1997;127:666–74.
22. Malone SM. Knowledge management: white knight or white elephant. Top Health Inf Manage 2001;21(3):33–43.
23. Petryschen P OBPL, Shamian J. Outcomes Monitoring: Adjusting for Risk Factors, Severity of Illness, and Complexity of Care. JAm Med Inform Assoc 1995;2:243–9.
24. Bond S TL. Issues in Measuring Outcomes in Nursing. J Adv Nurs. 1991;16(12):1492–1502.
25. Schneider EC, Riehl V, Courte-Wienecke S, Eddy DM, Sennett C. Enhancing performance measurement: NCQA's road map for a health information framework. National Committee for Quality Assurance. [see comment]. JAMASep 22–29 1999;282(12): 1184–90.
26. Duff L, Casey A. Implementing Clinical Guidelines: How can Informatics help? J Am Med Inform Assoc. May–Jun 1998;5(3): 225–6.

27. Lorenzi NM, Riley RT. Organizational issues=change. Int J Med Inform Mar 2003;69(2–3):197–203.

28. Swanson E. Management Information Systems: Appreciation and Involvement. Manag Sci. 1974;21:178–88.

29. Blackler F. Information systems design and planned organizational change: applying Unger's theory of social reconstruction. Special Issue: Methods and Frameworks for System Design. Behav Inform Technol 1992;11:175–83.

30. Blackler FOD. Information Technology & People: Designing for the Future. Cambridge, MA: MIT Press; 1987.

31. Sage AP. Behavioral and organizational considerations in the design of information systems and processes for planning and decision support. IEEE Trans Syst Man Cybernet. 1981;11:640–78.

32. Hoffman GM. The Technology Payoff: How to Profit with Empowered Workers in the Information Age. Burr Ridge, IL: Irwin Professional Publishing; 1994.

33. Davenport T. Saving IT's soul, human-centered information management. Harv Bus Rev. 1994;72:119–33.

34. Schultheiss E. Optimizing the Organization: How to Link People and Technology. Cambridge, MA: Ballinger Publishing Company; 1988.

35. Manklin D, Tora K, Barbara A. Factors in Successful Implementation of Computer-Based Office Information Systems: A Review of the Literature with Suggestions for OBM Research. Special Issue: Computers, People and Productivity. J Org Behav Manag 1984;6:1–20.

36. Churchman C, Schainblatt A. The Researcher and the Manager: a Dialectic of Implementation. Manag Sci. 1965;11:B–69–87.

37. Ives B, Olson M. User Involvement and MIS Success: A Review of Research. Manag Sci. 1984;30:586–603.

38. Lorenzi N, Riley R. Organizational Aspects of Health Informatics: Managing Technological Change. New York, NY: Springer Verlag; 1995.

39. Creswell JW. Research Design: Qualitative and Quantitative Approaches. Thousand Oaks, CA: SAGE Publications, Inc; 1994.

40. Dewitz S. System Analysis and Design, and the Transition to Objects. 1996.

41. Engh CA, Engh CA, Jr., Nagowski JP, Hopper RH, Jr. Database production and maintenance. Clin Orth Rel Res Apr 2004(421):35–42.

42. Zielstorff RD. Capturing and Using Clinical Outcome Data: Implications for Information Systems Design. J Am Med Inform Assoc. May–Jun 1995;2(3):191–6.

43. Barr JT. The Outcomes Movement and Health Status Measures. J All Health. 1995;24:13–28.

44. Brennan TA, Hebert LE, Laird NM, et al. Hospital characteristics associated with adverse events and substandard care. JAMA. Jun 26 1991;265(24):3265–69.

45. Petryshen P, Pallas LL, Shamian J. Outcomes monitoring: adjusting for risk factors, severity of illness, and complexity of care. J Am Med Inform Assoc Jul–Aug 1995;2(4):243–9.

46. Crane S. A Research Agenda for Outcomes Research. In: Patient Outcomes Research: Examining the Effectiveness of Nursing Practice. Paper presented at: State of the Science Conference, 1992; Bethesda, MD.

47. Last J. A Dictionary of Epidemiology, 3rd Edition. New York: Oxford University Press; 1995.

48. Dienemann J, Van de Castle B. The Impact of Healthcare Informatics on the Organization. J Nurs Admin Nov 2003; 33(11):557–62.

49. Nielsen J. Usability Engineering. Boston, MA: AP Professional; 1993.

50. Neely A, Sittig D. Basic Microbiologic and Infection Control Information to Reduce the Potential Transmission of Pathogens to Patients via Computer Hardware. J Am Med Inform Assoc 2002;9:500–8.

51. Niland J, Rouse L. Clinical Research Needs. In: Lehmann H, Abbott P, Roderer N, et al., eds. In: Aspects of Electronic Health Record-Systems (2nd ed). New York, London: Springer, 2006; 31–46.

52. Gregg AC. Performance management data systems for nursing service organizations. J Nurs Admin Feb 2002;32(2):71–8.

53. Dick RS, Steen EB, Detmer DE, Institute of Medicine (U.S.) Committee on Improving the Patient Record. The Computer-Based Patient: An Essential Technology for Health Care (Revised Ed.). Washington DC: National Academy Press, 1997.

54. https://cabig.nci.nih.gov/caBIG.

55. Cimino JJ, Clayton PD. Coping with changing controlled vocabularies. Proc Annu Symp Comp Appl Med Care 1994;135–9.

56. McCray AT, Srinivasan S, Browne AC. Lexical methods for managing variation in biomedical terminologies. Proc Annu Symp Comp Appl Med Care 1994:235–9.

57. Jean F MJ, Codaccioni A, Lavril M, Saquet D, Degoulet P. Using a meta-model to build a connection in an object oriented medical application development environment. Proc Am Med Inform Assoc Annu Symp 1994:483–7.

58. Campbell KE, Musen MA. Representation of clinical data using SNOMED III and conceptual graphs. Proc Annu Symp Comp Appl Med Care 1992:354–8.

59. Cote R, Rothwell D, Palotay J, Beckett R, L B. The Systemized Nomenclature of Human and Veterinary Medicine: SNOMED International. College of American Pathologists. 1993.

60. Spackman KA, Campbell KE. Compositional concept representation using SNOMED: towards further convergence of clinical terminologies. Proc AMIA. Annu Symp. 1998:740–4.

61. http://snomed.org.

62. http://www.cap.org/apps/cap.portal?_nfpb=true&_pageLabel= snomed_page.

63. http://www.loinc.org/background.

64. Hansen MT, Nohria N, Tierney T. What's your strategy for managing knowledge? Harv Bus Rev. 1999 Mar–Apr 1999;77(2): 106–16.

65. Marks RG CM, Ruberg SJ. Paradigm shifts in clinical trials enabled by information technology. Stat Med. 2001;20(17–18): 2683–96.

66. Ozbolt JG FJ, Hayden JR. Toward Data Standards for Clinical Nursing Information. J Am Med Inform Assoc. 1994;1:175–85.

67. Kobryn C. UML 2001: A Standardization Odyssey. Comm Assoc Comp Mach. 1999;42(10):29–37.

68. Jones J, Stahl D, Nichol M, Azen S. Enabling Clinically Based Knowledge Discovery in Pharmacy ClaHQIS Data: An Application in Bioinformatics. J Jap Soc Comp Stat. 2003;15:39–47.

69. Hewitt JB SS, Weber GI. Enterprise Identity Management Strategies. JHealthc Inform Manag. 2002;16(3):40–6.

70. Orfali R, Harkey D, Edwards J. Essential Client/Server Survival Guide, Second Edition. 2nd ed. New York: John Wiley & Sons Inc; 1996.

71. Ozbolt JG, Fruchtnicht JN, Hayden JR. Toward data standards for clinical nursing information. J Am Med Inform Assoc. Mar–Apr 1994;1(2):175–85.

72. Zielstorff R, Grobe S. National Commission on Nursing Implementation Project (NCNIP) Task Force on Nursing Information Systems. Next Generation of Nursing Information Systems: Essential Characteristics for Practice Professionals. Washington D.C.: American Nurses Publishing; 1993.

73. National Information Standards Organization (NISO). Understanding Metadata. Bethesda, MD: NISO Press; 2001.

74. Gray PWH. Decision Support in the Data Warehouse. Upper Saddle River, NJ: Prentice Hall Professional Technical Reference; 1998.

75. Fine LG, Keogh BE, Cretin S, Orlando M, Gould MM, Experience UKCS. How to evaluate and improve the quality and credibility of an outcomes database: validation and feedback

study on the UK Cardiac Surgery Experience. Br Med J. Jan 4 2003;326(7379):25–8.

76. Barton AJ, Russin MM. Re-engineering intensive care: the role of informatics. AACN Clin Iss. May 1997;8(2):253–61.

77. Benson Ar, Pregler J, Bean J, Rademaker A, Eshler B, Anderson K. Oncologists' Reluctance to Accrue Patients onto Clincal Trials: an Illinois Cancer Center Study. J Clin Oncol. 1991;9:2067–75.

78. Huber D, Schumacher L, Delaney C. Nursing Management Minimum Data Set (NMMDS). J Nurs Admin. 1997;27(4):42–8.

79. Badger KA. Patient care report cards: an analysis. Outcomes Management for Nursing Practice. Jan–Mar 1998;2(1):29–36.

80. Lenz S, Myers S, Nordlund S, Sullivan D, Vasista V. Benchmarking: finding ways to improve. Joint Comm J Qual Improv. May 1994;20(5):250–9.

81. Winn R, Botnick W, Dozier N. The NCCN Guideline Development Program. Oncology 1996;10(11S):23–8.

82. (AHIMA) AHIMA. Inventing the Future: Vision 2006. Chicago: AHIMA; 1996.

83. Edler L. Biometry—The Role of the Biostatician. Heidelberg: German Cancer Research Center; 2001.

84. Thall P. Ethical Issues in Biostatistics. Stat Meth Med Res. 2002;11(5):429–48.

85. Pocock S. Clinical Trials: A Practical Approach. New York: Wiley; 1987.

86. Suri S. Using Medical and Information Technology for Improving Quality of Care. J Healthc Inform Manag. 2002;25(1):33–9.

87. Deshpande AM, Brandt C, Nadkarni PM. Metadata-driven ad hoc query of patient data: meeting the needs of clinical studies. J Am Med Inform Assoc. Jul–Aug 2002;9(4):369–82.

88. Thomsen E. Solutions: Building Multidimensional Information Systems. New York: Wiley; 2000.

89. Mor V LL, Petrisek AC, Intrator O, Wachtel T, Maddock PG, Bland KI. Impact of breast cancer treatment guidelines on surgeon practice patterns: Results of a hospital-based intervention. Surgery. 2000;128(5):847–61.

90. Grobe SJ. Nursing Informatics 1997 postconference on patient guidelines and clinical practice guidelines: the state of our knowledge and a vision. J Am Med Inform Assoc. May–Jun 1998;5(3):315–6.

91. Collen M. A History of Medical Informatics in the United States, 1950–1990. Washington, D.C.: American Medical Informatics Association; 1995.

92. Eason K. Methodological Issues in the Study of Human Factors in Teleinformatic Systems. Behav Inform Tech. 1983;2(4):357–64.

93. Docherty J. Measuring Outcomes. In: Sederer L, ed. Outcomes Assessment in Clinical Practice. Philadelphia: Williams and Wilkins; 1996:8–18.

94. Bovbjerg VE, Olachanski V, Zimberg SE, Green JS, Rositer LF. Internet-based Monitoring and Benchmarking in Ambulatory Surgical Centers. Joint Comm J Qual Improv. 2000;26(8):450–65.

95. Klein E ea. Collection of data in clinical studies via the Internet. Studies Health Tech. Inform. 1997;43(Pt A):57–60.

96. Rind DM, Kohane I, Szolovits P, Safran C, Chueh HC, Barnett GO. Maintaining the confidentiality of medical records shared over the Internet and World Wide Web. Ann Intern Med. 1997;127(2):138–41.

97. Trinkus VPea. A Multi-institutional Internet-based hysterectomy database. Obstet Gynecol 2000;95(Suppl 1):S25.

98. Niland JC. NCCN outcomes research database: data collection via the Internet. Oncology (Huntington). Nov 2000;14(11A):100–3.

An Informatics Blueprint for Healthcare Quality Information Systems

Joyce C Niland, Layla Rouse and Douglas C Stahl

Updated information and services can be found at:

These include:

References

This article cites 66 articles, 11 of which can be accessed free at:
http://jamia.bmj.com/content/13/4/402.full.html#ref-list-1

Article cited in:
http://jamia.bmj.com/content/13/4/402.full.html#related-urls

Email alerting service

Receive free email alerts when new articles cite this article. Sign up in the box at the top right corner of the online article.

Notes