Quality Improvement Registries

Draft White Paper for Third Edition of

"Registries for Evaluating Patient Outcomes: A User's Guide"

Introduction

Quality assessment/ improvement registries (QI registries) seek to use systematic data collection and other tools to improve quality of care. While much of the information contained in the other chapters of this document applies to QI registries, these types of registries face unique challenges in the planning, design, and operation phases. The purpose of this paper is to describe the unique considerations related to QI registries.

As described in Chapter 1, a patient registry is largely defined by the population, exposure, outcomes of interest, and purpose. While a QI registry may have many purposes, at least one purpose is quality improvement. These registries generally fall into two categories: registries of patients exposed to particular health services (e.g., procedure registry, hospitalization registry) around a relatively short period of time (i.e., an event); and those with a disease/condition tracked over time through multiple provider encounters and/or multiple health services. An important commonality is that one exposure of interest is to health care providers/health care systems. These registries exist at the local, regional, national, and international levels.

QI registries are further distinguished from other types of registries by the tools that are used in conjunction with the systematic collection of data to improve quality at the population and individual patient levels. QI registries leverage the data about the individual patient or population to improve care in a large variety of ways. Examples of tools that facilitate data use for care improvement include patient lists, decision support (typically based on clinical practice guidelines), automated notifications, communications, and patient and population level reporting. For example, a diabetes registry managed by a single institution might provide a listing of all patients in a provider's practice that have diabetes and that are due for a clinical exam or other assessments. Decision support tools exist that read the structured data on the patient being provided to the registry and feedback recommendations for care based on evidence-based guidelines. This is a well-reported feature of the American Heart Association's Get With

The Guidelines® registries.[1] Certain registry tools will automatically notify a provider if the patient is due for a test, exam, or other milestone. Some tools will even send notifications directly to patients indicating that they are due for an action such as a flu shot. Reports are a key part of quality improvement. These range from reports on individual patients, such as a longitudinal report tracking a key patient outcome, to reports on the population under care by a provider or group of providers either alone or in comparison to others (at the local, regional, or national level). Examples of the latter reports include those that measure process of care (e.g., whether specific care was delivered to appropriate patients at the appropriate time) and those that measure outcomes of care (e.g., average Oswestry score results for patients undergoing particular spine procedures versus similar providers).

QI registries can further support improved quality of care by providing providers and their patients with more detailed information based on the aggregate experience of other patients in the registry. This can include both general information on the natural history of the disease process from the accumulated experience of other patients in the registry as well as more individual-patient level information on specific risk calculators that might help guide treatment decisions. Registries that produce patient-specific predictors of short and long-term outcomes (which can inform patients about themselves) as well as provider-specific outcomes benchmarked against national data (which can inform patients about the experience and outcomes of their providers) can be the basis of both transparent and shared decision-making between patients and their providers.

In addition to these examples are tools that are neither electronic nor necessarily provided through the registry systems. Non-electronic examples range from internal rounds to review registry results and make action plans, to quality-focused national or regional meetings that review treatment gaps identified from the registry data and teach solutions, to printed posters and cards or other reminders that display the key evidence-based recommendations that are measured in the registry. Further, even electronic tools need not be delivered through the registry systems themselves. While in many cases the registries do provide the functionality described above, it serves the same purpose if an electronic health record (EHR) provides access to decision support relevant to the goals of the patient registry. In other words, what characterizes QI registries is not the embedding of the tools in the registry but the use of the tools by the providers that participate in the registry to improve the care that they provide and the use of the registry to measure that improvement.

Planning

As described in Chapter 2 (Planning a Registry)[1], developing a registry starts with thoughtful planning and goal setting. Planning for a QI registry follows most of the steps outlined in Chapter 2, with some noteworthy differences and additions. A first step in planning is identifying key stakeholders. Similar to other types of registries, regional and national QI registries benefit from broad stakeholder representation, which is necessary but not sufficient for success. In QI registries, the provider needs to be engaged and active as the program is not simply supporting a surveillance function or providing a descriptive or analytic function but often focused on patient and/or provider behavior change. In many QI registries, these active providers are termed 'champions' and are vital for success, particularly early in development.[2] At the local level, the champions are typically the ones asking for the registry and almost by definition are engaged. Selecting stakeholders locally is generally focused on involving those with direct impact on care or those that can support the registry with information, systems, or labor. Yet, the common theme for both local and national QI registries is that the local champions must be successful in actively engaging their colleagues in order for the program to go beyond an 'early adopter' stage and to be sustainable within any local organization. Once a registry matures, other incentives may drive participation (e.g., recognition, competition, financial rewards, regulatory requirements), but the role of the champion in the early phases cannot be overstated.

Second, in order for a QI registry to meet its goal of improving care, it must provide actionable information for providers and/or participants to be able to modify their behaviors, processes, or systems of care. Actionable information can be provided in the form of patient outcomes measures (e.g., mortality, functional outcomes post discharge) or process of care or quality measures (e.g., compliance with clinical guidelines). While the ultimate goal of a QI registry is to improve patient outcomes by improving quality of care, it is not always possible for a QI registry to focus on patient outcomes measures. In some cases, outcome measures may not exist in the disease area of interest, or the measures may require data collection over a longer period than is feasible in the registry. As a result, QI registries have often focused on process of care or quality measures. While this has been criticized as less important than focusing on measures of patient outcomes, it should be noted that quality measures are generally developed from evidence-based guidelines and emphasize interventions that have been shown

[1] Chapters referenced in this document can be found in the second edition of "Registries for Evaluating Patient Outcomes: A User's Guide," available at: http://www.effectivehealthcare.ahrq.gov/ehc/products/74/531/Registries%202nd%20ed%20final%20to%20Eisenberg%209-15-10.pdf.

to improve long term outcomes, increasingly recognized through standardized processes (e.g., National Quality Forum), and are inherently actionable). Patient outcomes measures, on the other hand, do not yet have consensus across many conditions, are prone to bias in patients lost to follow-up, and may be expensive and difficult to collect reliably. Furthermore, long-term outcomes are generally not readily available for rapid cycle initiatives and may be too distant temporally from when the care is delivered to support effective behavior change. Despite these challenges, there has been an increasing focus in recent years on including outcome measures instead of or in addition to process of care measures in QI registries. This shift is driven in part by research documenting the lack of correlation between process measures and patient outcomes[3,4,5] and by arguments that health care value is best defined by patient outcomes, not processes of care.[6]

Selecting measures for QI registries typically requires balancing the goals to be relevant and actionable with the desire to meet other needs for providers, such as reporting quality measures to different parties (e.g., accreditation organizations, payers). Frequently, this is further complicated by the lack of harmonization between those measure requirements even in the same patient populations.[7] Even when there is agreement on the type of intervention to be measured and how the intervention is defined, there still may be variability in how the cases that populate the denominator are selected (e.g., by clinical diagnosis, by ICD-9 classification, by CPT codes). In the planning stages of a QI registry, it is useful to consider key parameters for selecting measures. The National Quality Forum offers the following four criteria for measure endorsement, which also apply to measure selection:

1) Important to measure and report to keep our focus on priority areas, where the evidence is highest that measurement can have a positive impact on healthcare quality.
2) Scientifically acceptable, so that the measure when implemented will produce consistent (reliable) and credible (valid) results about the quality of care.
3) Useable and relevant to ensure that intended users — consumers, purchasers, providers, and policy makers — can understand the results of the measure and are likely to find them useful for quality improvement and decision-making.
4) Feasible to collect with data that can be readily available for measurement and retrievable without undue burden.[8]

The National Priorities Partnership[9] and the Measure Applications Partnership,[10] both of which grew out of the National Quality Forum and provide support to the U.S. Department of Health and Human Services on issues related to quality initiatives and performance measurement, also offer useful criteria to consider when selecting measures.

One approach to consider in selecting measures is performing a cross-sectional assessment using the proposed panel of measures to identify the largest gaps between what is recommended in evidence-based guidelines or expected from the literature and what is actually done ('treatment gaps'). The early phase of the registry can then focus on those measures with the most significant gaps and for which there is a clear agreement among practicing physicians that the measure reflects appropriate care. The planning and development process should move from selecting measures to determining which data elements are needed to produce those measures (see the "Design" section). Measures should ideally be introduced with idealized populations of patients in the denominator for whom there is no debate about the appropriateness of the intervention. This may help reduce barriers to implementation that are due to physician resistance based on concerns about appropriateness in individual patients.

Once the measures and related data elements have been selected, pilot testing may be useful to assess the feasibility and burden of participation. Pilot testing may identify issues with the availability of some data elements, inconsistency in definition of data elements across sites, or barriers to participation, such as burden of collecting the data or disagreement about how exclusion criteria are constructed when put into practice. In order for the registry to be successful, participants must find the information provided by the registry useful for measuring and then modifying their behaviors, processes, or systems of care. Pilot testing may enable the registry to improve the content or delivery of reports or other tools prior to the large-scale launch of the program. If pilot testing is included in the plans for a QI registry, the timeline should allow for subsequent revisions to the registry based on the results of the pilot testing.

Change management is also an important consideration in planning a QI registry. QI registries need to be nimble in order to adapt to two continual sources of change. First, new evidence comes forward that changes the way care should be managed, and it is incumbent on the registry owner to make changes so that the registry is both current and relevant. In many registries, such as American Heart Association's Get With The Guidelines® Stroke program and the American Society of Clinical Oncologists' QOPI registry, this process occurs more than once per year. Second, registry participants manage what they measure, and, over time, measures can be rotated in or out of the panel so that attention is focused where it is most critical to overcome a continuing treatment gap or performance deficiency. This requires that the registry have standing governance to make changes over time, a system of data collection and reporting that is flexible enough to rapidly incorporate changes with minimal or no disruption to participants, and sufficient resources to communicate with and train participants on the changes. The governance structure should include individuals who are expert in the area of measurement science as well as in the scientific content. The registry system also needs to continuously respond to additional demands for transmitting quality measures to other parties that may or may not be harmonized (e.g.,

Physician Quality Reporting System, Meaningful Use reporting, Bridges to Excellence, state department of public health requirements). From a planning standpoint, QI registries should expect ongoing changes to the registry and plan for the resources required to support the changes. While this complicates the use of registry data for research purposes, it is vital that the registry always be perceived first as a tool for improving outcomes. Therefore, whenever changes are made to definitions, elements, or measures, these need to be carefully tracked so that analyses or external reporting of adherence may take these into account if they span time periods in which changes occurred.

Legal and Institutional Review Board Issues

As discussed in the legal/regulatory chapter, the new chapter on informed consent, and the new chapter on data protection, registries navigate a complex sea of legal and regulatory requirements depending on the status of the developer, the purpose of the registry, whether or not identifiable information is collected, the geographic locations in which the data are collected, and the geographic locations in which the data are stored (state laws, international laws, etc.). QI registries face unique challenges in that many institutions' legal departments and Institutional Review Boards (IRBs) may have less familiarity with registries for quality improvement, and, even for experts, the distinction between a quality improvement activity and research may be unclear.[11,12,13,14] Some research has shown that IRBs differ widely in how they differentiate research and quality improvement activities.[15] What is clear is that IRB review and, in particular, informed consent requirements, may not only add burden to the registry but may create biased enrollment that may in turn affect the veracity of the measures being reported.[16] Potential limitations of the IRB process have been identified in other reports, including for comparative effectiveness research, and will not be reviewed here.

For QI registries, which generally fit under the HIPAA health care operations definition, the issues that lead to complexity include whether or not the registry includes research as a primary purpose or any purpose, whether the institutions or practices fall under the Common Rule, and whether informed consent is needed. The Common Rule is discussed in the legal/regulatory chapter, and informed consent and quality improvement activities are discussed in the new chapter on informed consent. To assist in determining whether a quality improvement activity qualifies as research, the Office for Human Research Protections (OHRP) provides information in the form of a "Frequently Asked Questions" webpage.[17] OHRP notes most quality improvement activities are not considered research and therefore are not subject to the protection of human subjects regulations. However, some quality improvement activities are considered research, and the regulations do apply in those cases. To help determine if a quality

improvement activity constitutes research, OHRP suggests addressing the following four questions, in order:

> "(1) does the activity involve *research* (45 CFR 46.102(d));
> (2) does the research activity involve human subjects (45 CFR 46.102(f));
> (3) does the human subjects research qualify for an exemption (45 CFR 46.101(b)); and
> (4) is the non-exempt human subjects research conducted or supported by HHS or otherwise covered by an applicable FWA approved by OHRP."[18]

In addressing these questions, it is important to note the definition of "research" under 45 CFR 46.102(d). "Research" is defined as "…a systematic investigation, including research development, testing and evaluation, designed to develop or contribute to generalizable knowledge…." OHRP does not view many quality improvement activities as "research" under this definition and provides some examples of the types of activities that are not considered research.[19] It is also important to note the definition of "human subjects" under 45 CFR 46.101(b). "Human subject" is defined as "a living individual about whom an investigator (whether professional or student) conducting research obtains (1) Data through intervention or interaction with the individual, or (2) Identifiable private information." Again, OHRP does not view some quality improvement activities as collecting data on human subjects because data are not identifiable and were not collected through interaction with the individual patient (e.g., abstracted from a medical record).[20]

These questions provide some helpful information in determining whether a quality improvement registry is subject to the protection of human subjects regulations, but some researchers and IRBs have still reported difficulty in this area.[21,22] Remaining questions include, for example, if the registry includes multiple sites, is separate IRB approval from every institution required? If the registry is considered research, in what circumstances is informed consent required?

There have been several recent calls to refine and streamline the IRB process for QI registries,[23] and some of this work is advancing. Recently, OHRP has proposed revisions to the Common Rule that would address some of these issues; the proposed changes were posted for a public comment period, which closed in October 2011.[24] Without some changes and greater clarity around existing regulations as they relate to QI registries, it will be difficult for some registries to be successful.

Design

Designing a quality improvement registry presents several challenges, particularly when multiple stakeholders are involved. Staying focused on the registry's key purposes, limiting respondent burden, and being able to make use of all of the data collected are practical considerations in developing programs. First, the type of quality improvement registry needs to be determined. Is the goal to improve the quality of patients with a disease or patients presenting for a singular event in the course of their disease? For example, a QI registry in cardiovascular disease will be different (i.e. sampling, endpoints, measures) if it focuses on patients with coronary artery disease versus if it focuses on patients with a hospitalization for acute coronary syndrome or patients who undergo percutaneous coronary angioplasty as an inpatient or outpatient. In the first example, the registry may need to track patients over time and across different providers; reminder tools may be needed to prompt follow-up visits or lab tests. In the second example, the registry may need to collect detailed data at a single point in time on a large volume of patients.

Second, QI registries that collect data within a single institution differ from those that collect data at multiple institutions regionally or nationally. Single institution registries, for example, may be designed to fit within specific workflows at the institution or to integrate with one EHR system. They may reflect the specific needs of that institution in terms of addressing treatment gaps, and they may be able to obtain participant buy-in for reporting plans (e.g., for unblinded reporting). Regional or national level registries, on the other hand, must be developed to fit seamlessly into multiple different workflows. These registries must address common treatment gaps that will be relevant to many institutions, and they must develop approaches to reporting that are acceptable to all participants.

The appropriate level of analysis and reporting is an important consideration for designers of QI registries. Reports may provide data at the individual patient, provider, or institution level, or they may provide aggregate data on groups of patients, providers, and institutions. The aggregate groups may be based on similar characteristics (e.g., disease state, hospitals of a similar size), geography, or other factors. The registry may also provide reports to the registry participants, to patients, or to the public. Reports may be unblinded (e.g., the provider is identifiable) or blinded, and they may be provided through the registry or through other means. In designing the registry, consideration should be given to what types of reports will be most relevant for achieving the registry's goals, what types of reports will be acceptable to participants, and how those reports should be presented and delivered. Reporting considerations are discussed further in the "Reporting to Providers and the Public" section.

As described above, there are many challenges in selecting existing measures or designing and testing new measures. Once measures have been selected, the 'core data' can be determined. Since QI registries are part of health care operations, it is critical that they do not overly interfere with the efficiency of those operations, and therefore the data collection must be limited to those data elements that are essential for achieving the registry's purpose. One approach to establishing the core data set is to first identify the outcomes or measures of interest and then work backwards to the minimal data set, adding those elements required for risk adjustment or relevant subgroup analyses. For example, the inclusion and exclusion criteria for a measure, as well as information used to group patients into numerator and denominator groups, can be translated into data elements for the registry. The "Using Performance Measures to Develop a Dataset" case example describes this process for the Get With The Guidelines® Stroke program. Depending on the goals of the registry, the core data set may also need to align with data collection requirements for other quality reporting programs.

Many QI registries have gone further by establishing a core data set and an enhanced data set for participating groups that are ready to extend the range of their measurements. This tiered model can be very effective in appealing to a broad range of practices or institutions. Examples include the Get With The Guidelines® program, which allows hospitals to select performance measures or both performance and quality measures, and the American College of Surgeons NSQIP program, which has a core data set and the ability to add targeted procedure modules.

QI registries also may need to develop sampling strategies during the design phase. The goal of sampling in quality improvement registries is to provide representativeness (i.e., reflective of the patients treated by the physician or practice) and precision (i.e., sufficient sample size to provide reasonable intervals around the metrics generated from each practitioner/practice to be useful in before/after or benchmarking comparisons). Sampling frames need to balance simplicity with sustainability. For example, an all comers model is easy to implement but can be difficult to sustain, particularly if the registry utilizes longitudinal follow-up. For example, an orthopedic registry maintained by a major U.S. center sought to enroll all patients presenting for hip and knee procedures. Since the center performed several thousand procedures each year, within a few short years the numbers of follow-ups being performed climbed to the tens of thousands. This was both expensive and likely unsustainable. On the other hand, a sampling frame can be difficult and confusing. While a sampling frame can be readily administered in a retrospective chart review, it is much more difficult to do so in a prospective registry. Some approaches to this issue have included selecting specific days or weeks in a month for patient enrollment. But, if these frames are known to the practitioners, they can be 'gamed,' and auditing may be necessary to determine if there are sampling inconsistencies. Pilot testing can be useful for assessing the pace of

patient enrollment and the feasibility of the sampling frame. Ongoing assessments may also be needed to ensure that the sampling frame is yielding a representative population.

An additional implication when considering how to implement a sampling strategy is that for QI registries in which concurrent case ascertainment and intervention is involved, only those patients that are sampled may benefit from real-time QI intervention and decision-support. In these circumstances, patients who are not sampled are also less likely to receive the best care. This disparity may only increase as EHR-enabled decision support becomes increasingly sophisticated and commonplace.

Operational Considerations

As with most registries, the major cost for participants in a QI registry is data collection and entry rather than the cost of the data entry platform or participation fees. Because QI registries are designed to fit within existing health care operations, many of the data elements collected in these registries are already being collected for other purposes (e.g., claims, medical records, other quality reporting programs). QI registries are often managed by clinical staff who are less familiar with clinical research and who must fit registry data collection into their daily routines. Both of these factors make integration with existing health information technology systems or other data collection programs attractive options for some QI registries. Integration may take many forms. For example, data from billing systems may be extracted to assist with identifying patients or to pull in basic information on the patients. EHRs may contain a large amount of the data needed for the registry, and integration with the EHR system could substantially reduce the data collection burden on sites. However, integration with EHRs can be complex, particularly for registries at the regional or national level that need to extract data from multiple systems. A critical challenge is that the attribution of clinical diagnoses in the context of routine patient care is often not consistent with the strict coding criteria for registries, making integration with EHR systems more complex. Chapter 11 discusses integration of registries with EHR systems. Another alternative for some disease areas is to integrate data collection for the registry with data collection for other quality initiatives (e.g., Joint Commission, CMS). Typically, these types of integration can only provide some of the necessary data; participants must collect and enter additional data to complete the CRFs.

The burden of data collection is an important factor in participant recruitment and retention. Much of the recruitment and retention discussion in Chapter 9 (Recruitment and Retention of Patients and Providers) applies to QI registries. However, one area in which QI registries differ from other types of registries is in the motivations for participation. Sites may participate in other registries because of interest in the research question or as part of mandated participation for state or federal payment or regulatory

requirements. When participation is for research purposes, they may hope to connect with other providers treating similar patients or contribute to knowledge in this area. In contrast, participants in QI registries expect to use the registry data and tools to effect change within their organization. Participation in a QI registry and related improvement activities can require significant time and resources, and incentives for participation must be tailored to the needs of the participants. For example, recognition programs, support for QI activities, QI tools, and benchmarking reports may all be attractive incentives for participants. In addition, tiered programs, as noted above, can be an effective approach to encouraging participation from a wide variety of practice or institution types. Understanding the clinical background of the stakeholders (e.g., nurses, physicians, allied health, and quality improvement professionals) and their interest in the program is critical to designing appropriate incentives for participation.

Quality Improvement Tools

As described above, QI tools are a unique and central component of QI registries. QI tools generally leverage the data in the registry to provide information to participants with the goal of improving quality of care. Examples of QI tools that draw on registry data include patient lists, automated notifications and other types of communications, decision support tools, and reports. Generally, QI tools are designed to meet one of two goals: care delivery and coordination or population measurement. Care delivery and coordination tools aim to improve care at the individual patient level. For example, an automated notification may inform a provider that a specific patient is due for an exam. Population measurement tools track activity at the population level, with the goal of assessing overall quality improvement and identifying areas for future improvement activities. For example, a report may be used to track an institution's performance on key measures over time and compared to other similar institutions. These types of reports can be used to demonstrate both initial and sustained improvements. Table 1 below summarizes some common types of QI tools in these two categories and describes their uses.

Table 1: Common Quality Improvement Tools

Major Goal	QI Tool	Description
Care delivery and coordination	Patient lists	Lists of patients with a particular condition who may be due for an exam, procedure, etc.
	Patient level reports	Summarize data on an individual patient (e.g., longitudinal data on blood pressure readings).
	Automated notifications	Prompt provider or patient when an exam or other action is needed.
	Automated communications	Summarize patient information in a format that can be shared with the patient or other providers.
	Decision support	Provide recommendations for care for an individual patient

	tools	using evidence-based guidelines.
Population measurement	Population level standardized reports	Provide an analysis of population-level compliance with QI measures or other summaries (e.g., patient outcomes across the population).
	Benchmarking reports	Compare population-level data for various types of providers.
	Ad-hoc reports	Enable participants to analyze registry data to explore their own questions.
	Population level dashboards	Provide snapshot look at QI progress and areas for continued improvement.
	3rd party quality reporting	Enables registry data to be leveraged for reporting to 3rd party quality reporting initiatives.

QI registries may incorporate various tools, depending on the needs of their participants and the goals of the registry. Table 2 below describes the types of functionalities that have been implemented in three different registries – two at the national level and one at the regional level.

Table 2: Quality Improvement Tools Implemented in Three Registries

Registry	Disease/Condition	Functionalities Implemented
AHA Get With The Guidelines®	Heart failure Stroke	• Decision support (guidelines) • Communication tools • Patient education materials • Real-time quality reports with benchmarks • Transmission to 3rd parties
MaineHealth Clinical Improvement Registry	Diabetes	• Patient care 'gap' reports • Decision support • Transmission to 3rd parties
National Comprehensive Cancer Network (NCCN)	Cancer	• Patient care 'gap' reports • Center level reports • Education materials

Quality Assurance

In addition to developing data elements and QI tools, QI registries must pay careful attention to quality assurance issues. Quality assurance, which is covered in Chapter 10 (Data Collection and Quality Assurance), is important for any registry to ensure that appropriate patients are being enrolled and the data being collected are accurate. Data quality issues in registries may result from inadequate training, incomplete case identification or sampling, misunderstanding or misapplication of inclusion/exclusion criteria, or misinterpretation of data elements. Quality assurance activities can help to identify these types of issues and improve the overall quality of the registry data. QI registries can use quality assurance

activities to address these common issues, but they must also be alert to data quality issues that are unique to QI registries. Unlike other registries, many QI registries are linked to economic incentives, such as licensure or access to patients, incentive payments, or recognition or certification. These are strong motivators for participation in the registry, but they may also lead to issues with data quality. In particular, 'cherry picking,' which refers to the non-random selection of patients so that those patients with the best outcomes are enrolled in the registry, is a concern for QI registries. Whenever data are being abstracted from source documents by hand and then entered manually into electronic data entry systems, there is a risk of typographical errors, errors in unit conversions (e.g., 12 hour to military time, milligrams to grams). Automated systems for error checking can reduce the risk of errors being entered into the registry when range checks and valid data formats are built into the data capture platform.

Auditing is one approach to quality assurance for QI registries. Auditing may involve on-site audits, in which a trained individual reviews registry data against source documents, or remote audits, in which the source documents are sent to a central location for review against the registry data. Because auditing all sites and all patients is cost-prohibitive, registries may audit a percentage of sites and/or a percentage of patients. QI registries should determine if they will audit data, and, if so, how they will conduct the audits. A risk-based approach may be useful for developing an auditing plan. In a risk-based approach, the registry assesses the risk for intentional error in data entry or patient selection. Registries that may have an increased risk of intentional error are mandatory registries, registries with public reporting, or registries that are linked to economic incentives. Registries with an increased risk may decide to pursue more rigorous auditing programs than registries with a lower risk. For example, a voluntary registry with confidential reporting may elect to do a remote audit of a small percentage of sites and patients each year. A registry with public reporting that is linked to patient access, on the other hand, may audit a larger number of sites and patients each year, with a particular focus on key outcomes that are included in the publically reported measures.

Questions to consider when developing a quality assurance plan involving auditing include: what percentage of sites should be audited each year; what percentage of data should be audited (all data elements for a sample of patients or only key data elements for performance measures); how sites should be selected for auditing (random, targeted, etc.); on-site audits vs. remote audits; and what constitutes passing an audit. Depending on the purpose of the registry, quality assurance plans may also address issues with missing data (e.g., what percentage of missing data is expected? Are data missing at random?) or patients who are lost to follow-up (e.g., what lost to follow-up rate is anticipated? Are certain subgroups of patients more likely to be lost to follow-up?). Lastly, quality assurance plans must consider how to address data quality issues. Audits and other quality assurance activities may identify

problem areas in the registry data set. In some cases, such as when the problem is isolated to one or two sites, additional training may resolve the issue. In other cases, such as when the issue is occurring at multiple sites, data elements, documentation, or study procedures may need to be modified. In rare instances, quality assurance activities may identify significant performance issues at an individual site. The issues could be intentional (e.g., cherry picking) or unintentional (e.g., data entry errors). The registry should have a plan in place for addressing these types of issues.

Analytical Considerations

While registries are powerful tools for understanding and improving quality of care, several analytical issues need to be considered. In general, the observational design of registries requires careful consideration of potential sources of bias and confounding that exist due to the non-randomization of treatments or other sources. These sources of bias and confounding can threaten the validity of findings. Fortunately, the problems associated with observational study designs are well known, and a number of analytical strategies are available for producing robust analyses. Despite the many tools to handle analytical problems, limitations due to observational design, structure of data, measured and unmeasured confounding, and missing data should be readily acknowledged. Below is a brief description of several considerations when analyzing QI registry data and how investigators commonly address the problems.

Observational designs used in registries offer the ability to study large cohorts of patients, allowing for careful description of patterns of care or variations in practice compared to what is considered appropriate or best care. While not an explicit intention, registries are often used to evaluate an effect of a treatment or intervention. The lack of randomization in registries, which limits causal inferences, is an important consideration. For example, in a randomized trial, a treatment or intervention can be evaluated for efficacy because different treatment options have an equal chance of being assigned. Another important characteristic that observational studies may lack is the chance of actually receiving a treatment. In a randomized trial, subjects meet a set if inclusion criteria and therefore have an equal chance of receiving a given treatment. However, in a registry, there are likely patients that have no chance of receiving a treatment. As a result, some inferences cannot be generalized across all patients in the registry.

An inherent but commonly ignored issue is the structure of health or registry data. Namely, physicians manage patients with routine processes, and physicians practice within hospitals or other settings that also share directly or indirectly common approaches. These clusters or “hierarchical” relationships within the data may influence results if ignored. For example, for a given hospital, a type of procedure may be preferred due to similar training experiences from surgeons. Common processes or patient selections are

also more likely within a hospital than compared to another hospital. These observations form a cluster and cannot be assumed to be independent. Without accounting for the clustering of care, incorrect conclusions could be made. Models that deal with these types of clustered data, often referred to as hierarchical models, can address this problem. These models may also be described as multi-level, mixed, or random effects models. The exact approach depends on the main goal of an analysis, but typically includes fixed effects, which have limited number of possible values, and random effects, which represent a sample of elements drawn from a larger population of effects. Thus, a multilevel analysis allows incorporation of variables measured at different levels of the hierarchy and accounts for the attribute that outcomes of different patients under the care of a single physician or within the same hospital are correlated.

Adequate sample size for research questions is also an important consideration. In general, registries allow large cohorts of patients to be enrolled, but, depending on the question, sample sizes may be highly restricted (e.g., in the case of extremely rare exposures or outcomes). For example, a comparative effectiveness research question may address anticoagulation in patients with atrial fibrillation. As the analysis population is defined based on eligibility criteria, including whether patients are naïve to the therapy of interest, sample sizes with the exposure may become extremely small. Likewise, an outcome of angioedema may be extremely rare, and, if being evaluated with a new therapeutic, both the exposure and outcome may be too small of sample to fully evaluate. Thus, careful attention to the likely exposure population after establishing eligibility criteria as well as the likely number of events or outcomes of interest is extremely important. In cases where sample sizes become small, it is important to determine whether adequate power exists to reject the null hypothesis.

Confounding is a frequent challenge for observational studies, and a variety of analytical techniques can be employed to account for this problem. When a characteristic correlates with both the exposure of interest and the outcome of interest, it is important to account for the relationship. For example, age is often related to mortality and may also be related to use of a given process. In a sufficiently large clinical trial, age generally is balanced between those with and without the exposure or intervention. However, in an observational study, the confounding factor of age needs to be addressed through risk adjustment. Most studies will use regression models to account for observed confounders and adjust for outcome comparisons. Others may use matching or stratification techniques to adjust for the imbalance in important characteristics associated with the outcome. Finally, another approach being used more frequently is the use of propensity scores that take a set of confounders and reduce them into a single balancing score that can be used to compare outcomes within different groups.

As QI registries have evolved, an important attribute is defining eligibility for a process measure. The denominator for patients eligible for a process measure should be carefully defined based on clinical criteria, with those with a contraindication for a process excluded. The definition of eligibility in a process measure is critical for accurate profiling of hospitals and health care providers. Without such careful, clear definitions, it would be challenging to benchmark sites by performance.

With any registry or research study, data completeness needs to be considered when assessing the quality of the study. Reasons for missing data vary depending on the study or data collection efforts. For many registries, data completeness depends on what is routinely available in the medical record. Missing data may be considered ignorable if the characteristics associated with the missingness are already observable and therefore included in analysis. Other missing data may not ignorable either because of its importance or because the missingness cannot be explained by other characteristics. In these cases, methods for addressing the missingness need to be considered. Various options for handling the degree of missing data including discarding data, using data conveniently available, or imputing data with either simple methods (i.e., mean) or through multiple imputation methods.

Reporting to Providers and the Public

An important component of quality improvement registries is the reporting of information to participants, and, in some cases, to the public. The relatively recent origin of clinical data registries was directly related to early public reporting initiatives by the federal government. Shortly after the 1986 publication of unadjusted mortality rates by the Health Care Financing Administration, the predecessor of CMS, a number of states (e.g., the New York Cardiac Surgery Reporting System),[25,26] regions (e.g., Northern New England Cardiovascular Disease Study Group, or NNE),[27,28] government agencies (e.g., the Veteran's Administration),[29,30,31] and professional organizations (e.g., Society of Thoracic Surgeons)[32,33,34] developed clinical data registries. Many of these focused on cardiac surgery. Its index procedure, coronary artery bypass grafting (CABG) is the most frequently performed of all major operations, it is expensive, and it has well-defined adverse endpoints.

Registry developers recognized that the HCFA initiative had ushered in a new era of healthcare transparency and accountability. However, its methodology did not accurately characterize provider performance because it used claims data and failed to adjust for preoperative patient severity.[35] Clinical registries, and the risk-adjusted analyses derived from them, were designed to address these deficiencies. States such as New York, Pennsylvania, New Jersey, California, and Massachusetts developed public report cards for consumers, while professional organizations and regional collaborations used registry

data to confidentially feed back results to providers and to develop evidence-based best practice initiatives.[36,37]

The impact of public reporting on healthcare quality remains uncertain. One randomized trial demonstrated that heart attack survival improved with public reporting,[38] and there is evidence that low-performing hospitals are more likely to initiate quality improvement initiatives in a public reporting environment.[39] However, a comprehensive review[40] found generally weak evidence for the association between public reporting and quality improvement, with the possible exception of cardiac surgery, where results improved significantly after the initial publication of report cards in New York two decades ago.[41,42,43] Some studies have questioned whether this improvement was the direct result of public reporting, as contiguous areas without public reporting also experienced declining mortality rates.[44] Similar improvements have been achieved with completely confidential feedback or regional collaboration in northern New England[45] and in Ontario.[46] Thus, there appear to be many effective ways to improve healthcare quality—public reporting, confidential provider feedback, professional collaborations, state regulatory oversight—but the common denominator among them is a formal system for collecting and analyzing accurate, credible data,[47] such as registries.

Public reporting should theoretically affect consumer choice of providers and redirect market share to higher performers. However, empirical data failed to demonstrate this following the HCFA hospital mortality rate publications,[48] and CABG report cards had no substantial effect on referral patterns or market share of high and low performing hospitals in New York[49,50] or Pennsylvania.[51,52] Studies suggest numerous explanations for these findings, including lack of consumer awareness of and access to report cards; the multiplicity of report cards; difficulty in interpreting performance reports; credibility concerns; small differences among providers; lack of "newsworthiness"; the difficulty of using report cards for urgent or emergent situations; and the finite ability of highly ranked providers to accept increased demand.[53,54,55] Professor Judith Hibbard and colleagues have suggested report card formats that enhance the ability of consumers to accurately interpret accurate report cards, including visual aids (e.g., star ratings) that synthesize complex information into easily understandable signals.[56,57] A recent Kaiser Family Foundation survey[58] suggests that, particularly among more educated patients, the use of objective ratings to choose providers has steadily increased over the past decade, and health reform is likely to accelerate this trend.

The potential benefits of public reporting must be weighed against the unintended negative consequences, such as "gaming" of the reporting system.[59,60] The most concerning negative consequence is risk aversion, the reluctance of physicians and surgeons to accept high-risk patients because of their

anticipated negative effect on their report card ratings. Because these highest risk patients may derive the greatest benefit from aggressive intervention, risk aversion may produce a net decrement in public health and a net increase in long-term costs because the best treatments were not initially used.[61,62,63] Risk aversion unquestionably exists, but its extent and overall population impact are difficult to quantify. CABG risk aversion may have occurred in New York[64,65] and Pennsylvania,[66] but studies in California[67] and England[68] have not demonstrated similar findings. Numerous studies document probable risk aversion in percutaneous coronary interventions.[69,70,71] Possible approaches to mitigate risk aversion include demonstrating to providers the adequacy of risk-adjustment and modifying those models when appropriate; excluding highest risk patients from reporting; separate reporting of highest risk patients; and careful clinical review of patients turned down for interventions.

Irrespective of its end results, many believe that public reporting is a fundamental ethical obligation of physicians.[72,73] It addresses the patient right of autonomy or self-determination in decision-making. Whether or not they choose to exercise this right, patients making a choice about treatments should be fully informed, which arguably includes their right to know the comparative performance of potential providers.

When a decision has been made to publicly report outcomes, such measures must meet strict criteria. Professional organizations have emphasized the need to use high quality, audited clinical data whenever possible, and to employ the most appropriate statistical methodologies.[74,75] Professional society guidelines provide recommendations of varying strength and evidence strength, whereas performance measures should be a select subset of these guidelines that have the highest level of evidence and strongest class of recommendation (e.g., ACC/AHA class 1[recommended] or 3 [not indicated, or harmful], level A evidence). National Quality Forum requirements for performance measure endorsement have recently been updated. In addition to the four basic requirements of Importance, Scientific Acceptability, Usability, and Feasibility, they emphasize the need for robust, systematic evaluation of the evidence base and comprehensive testing of reliability and validity.[76,77,78]

The unit of analysis in public reporting is controversial. Many states report results for some procedures at the physician or surgeon level, but in many healthcare areas sample sizes and the small amount of variation attributable to the physician make it difficult to reliably discriminate performance.[79,80,81] Compiling data from a variety of process and outcome endpoints may help to mitigate sample size issues, as may aggregation of results over multiple years.

Report cards at the individual physician level may be more likely to cause risk aversion compared with group or hospital-level reports. Changes in health care delivery models must also be considered. As patient care is increasingly provided by teams of providers that may even cross traditional specialty boundaries, individual physician reporting may become less relevant and feasible. Reimbursement will increasingly be based on the overall care provided to a patient or population, and leaders will have a direct financial incentive to assess the performance of individual physicians in such care groups (e.g., ACOs), whether or not such results are publicly reported.

Use of QI Registry Data for Research Studies

An emerging trend is the use of data from QI registries to support additional studies. QI registries may collect large volumes of clinical data that can be used to support research studies. Studies using data from QI registries generally occur in one of two ways. First, the registry may be modified to collect additional data for a substudy. For example, a registry may collect in-hospital data on patients admitted to the hospital for a specific procedure. To study long-term outcomes of the procedure, the registry protocol may be modified to collect follow-up data for a subset of patients. An example of this approach was the OPTIMIZE-HF registry, which collected in-hospital data on patients admitted with heart failure. A subset of patients provided consent to be contacted after six months to collect additional data.[3] QI registries can also be modified to support other types of studies, such as studies where a subset of participating sites are randomized (cluster randomization) or a subset of patients are randomized (experimental trial). When modifying the registry protocol to support a substudy, the impact on the primary purpose of the registry must be considered, as well as any additional ethical or regulatory requirements introduced by the new data collection effort.

A second approach to using QI registries to support additional studies is to use the registry data, either alone or linked to another dataset. For example, a registry that collects in-hospital data may be linked to a claims database to obtain information on long-term outcomes or to examine other questions.[82] In these cases, the technical, legal, and ethical considerations related to linking registry datasets discussed in Chapter 7 should be reviewed. Regardless of which approach is used, researchers using data from a QI registry for additional research studies must understand how the data are collected and how patients are enrolled in the primary registry in order to draw appropriate conclusions from the new study.

Limitations of Current QI Registries

To summarize some of the key points above, the ideal QI registry collects uniform data on risk factors, treatments, and outcomes at key points for a particular disease or treatment. It obtains the data from multiple sources, across care settings and leverages existing health information technology (HIT) systems through interoperability and other data sets (from registries, claims, national indices, etc.) through linkage. Such a registry uses standardized methods to assure that the patients sampled are representative, that data are of high quality and that it is comparable across providers. Such registries provide feedback at the patient and population level, and, in addition to facilitating quality improvement, they perform quality reporting to third parties. Importantly, they maintain high levels of participation by providers and patients and have a long term, sustainable business model.

Clearly, most QI registries do not achieve the ideal. The term 'QI registries' is currently used to refer to a broad spectrum of registries, from local or regional registries aimed at improving care for a specific patient population to large, national registries with sophisticated benchmarking data. Many current QI registries focus on isolated conditions or procedures (e.g., the ACC NCDR Cath/PCI Registry; the STS Adult Cardiac Surgery Database). Health reform will require the acquisition of data about the overall, comprehensive care of conditions such as coronary artery disease, or of populations.[83] This may be facilitated by linkages among related data registries, which might include outpatient preventative care, inpatient acute care and procedures, rehabilitation, and chronic disease management.

Current QI registries also have temporal limitations. They characteristically collect data only in-hospital or for 30 days after admission or a procedure. However, patients, payers, and regulators are also interested in longer-term, longitudinal outcomes such as survival, readmission, reintervention, and cumulative resource use. Such information is useful for shared decision-making and for comparative effectiveness research. By linking together robust clinical data registries and administrative databases such as MEDPAR or the Social Security Death Master File[84,85] that provide long-term data, many of these current limitations of clinical registries would be mitigated.

In order for such linkages to be implemented, a number of challenges would need to be overcome. These include a lack of standardized data sets; difficulties collecting data across care settings; inability to leverage existing HIT systems to reduce duplication of clinician effort; inability to link to other data sources that might reduce data collection burden or enrich outcomes; significant variation in the quality of methods used to collect and report data; and quite different levels of participation and business models. Even registries in related conditions may not be fully compatible.

Potential solutions to such issues have been identified.[86] These include, for example, condition and cross-condition efforts to standardize common or core data element specifications, data quality and audit standards, and methodological considerations such as risk-adjustment. Collecting data across care settings will be improved by solving the patient identity management issues (discussed in the new patient identity management chapter), which will require clarification and perhaps revision of HIPAA and Common Rule regulations. Overcoming interoperability issues through the promulgation of open standards (e.g., HITSP TP-50) as described in Chapter 11 could have dramatic impact if adopted widely by EHR systems and registries.

Significant hospital data collection costs are additional limitations of clinical registries. Some data elements such as lab values may be automatically extracted from EHRs, but detailed clinical data may still require manual extraction. Existing national registries must develop sustainable business models, and there must be incentives and assistance for the development of new registries where none currently exist.

Conclusions

QI registries have documented success at improving quality of care at the local, regional, and national levels. While QI registries differ in their area of focus, choice of measures, and level of reporting, their consistent features are the use of systematic data collection and other tools to improve quality of care. QI registries also differ from other types of registries in many ways, such as in their use of provider "champions," the inclusion of actionable measures, the frequency of major changes to the registry data collection, the motivations for participation, and the use of blinded or unblinded quality reports to providers, and, in some cases, the public. Because of these differences, QI registries face special challenges, particularly in the planning, design, and operations phases. This paper describes those challenges and discusses best practices, where appropriate.

References

[1] Labresh KA, Gliklich R, Liljestrand J, et al. Using "Get With The Guidelines" to improve cardiovascular secondary prevention. Jt Comm J Qual Patient Safety 2003 Oct;29(10):539-50.
[2] Raval MV, Bentrem DJ, Eskandari MK, Ingraham AM, Hall BL, Randolph B, Ko CY, Morton JM. The role of Surgical Champions in the American College of Surgeons National Surgical Quality Improvement Program--a national survey. J Surg Res. 2011 Mar;166(1):e15-25. Epub 2010 Nov 25.
[3] Fonarow GC, Abraham WT, Albert NM, Stough WG, Gheorghiade M, Greenberg BH, O'Connor CM, Pieper K, Sun JL, Yancy C, Young JB; OPTIMIZE-HF Investigators and Hospitals. Association between performance measures and clinical outcomes for patients hospitalized with heart failure. JAMA. 2007 Jan 3;297(1):61-70.
[4] Lee JS, Primack BA, Mor MK, Stone RA, Obrosky DS, Yealy DM, Fine MJ. Processes of Care and Outcomes for Community-Acquired Pneumonia. Am J Med. 2011 Oct 13.
[5] Morse RB, Hall M, Fieldston ES, McGwire G, Anspacher M, Sills MR, Williams K, Oyemwense N, Mann KJ, Simon HK, Shah SS. Hospital-level compliance with asthma care quality measures at children's hospitals and subsequent asthma-related outcomes. JAMA. 2011 Oct 5;306(13):1454-60.
[6] Porter, ME. What Is Value in Health Care? N Engl J Med 2010; 363:2477-2481.
[7] Performance Measurement: Accelerating Improvement. Committee on Redesigning Health Insurance Performance Measures, Payment, and Performance Improvement Programs. Institute of Medicine. 2006. http://iom.edu/Reports/2005/Performance-Measurement-Accelerating-Improvement.aspx. Accessed October 30, 2011.
[8] "What NQF Endorsement Means." National Quality Forum. http://www.qualityforum.org/Measuring_Performance/ABCs/What_NQF_Endorsement_Means.aspx. Accessed October 30, 2011.
[9] National Priorities Partnership. http://www.nationalprioritiespartnership.org/Home.aspx. Accessed October 30, 2011.
[10] Measure Applications Partnership. National Quality Forum. http://www.qualityforum.org/Setting_Priorities/Partnership/Measure_Applications_Partnership.aspx. Accessed October 30, 2011.
[11] Casarett D, Karlawish JH, Sugarman J. Determining when quality improvement initiatives should be considered research: proposed criteria and potential implications. JAMA. 2000 May 3;283(17):2275-80.
[12] Dokholyan RS, Muhlbaier LH, Falletta JM, Jacobs JP, Shahian D, Haan CK, Peterson ED. Regulatory and ethical considerations for linking clinical and administrative databases. Am Heart J. 2009 Jun;157(6):971-82.
[13] Lynn J et al. The ethics of using quality improvement methods in health care. Ann Intern Med 146(9):666–673.
[14] Nerenz DR. Ethical issues in using data from quality management programs. Eur Spine J. 2009 Aug;18 Suppl 3:321-30. Epub 2009 Apr 14.
[15] Johnson N, Vermeulen L, Smith KM (2006) A survey of academic medical centers to distinguish between quality improvement and research activities. Qual Manag Health Care 15(4):215–220.
[16] Tu JV, Willison DJ, Silver FL, et al. Impracticability of informed consent in the registry of the Canadian stroke network. N Engl J Med. 2004;350:1414–21.
[17] "Quality Improvement Activities – FAQs." Office for Human Research Protections. U.S. Department of Health and Human Services. http://answers.hhs.gov/ohrp/categories/1569. Accessed August 24, 2011.
[18] How does HHS view quality improvement activities in relation to the regulations for human research subject protections? "Quality Improvement Activities – FAQs." Office for Human Research Protections. U.S. Department of Health and Human Services. http://answers.hhs.gov/ohrp/categories/1569. Accessed August 24, 2011.
[19] Do the HHS regulations for the protection of human subjects in research (45 CFR part 46) apply to quality improvement activities conducted by one or more institutions whose purposes are limited to: (a) implementing a

practice to improve the quality of patient care, and (b) collecting patient or provider data regarding the implementation of the practice for clinical, practical, or administrative purposes? "Quality Improvement Activities – FAQs." Office for Human Research Protections. U.S. Department of Health and Human Services. http://answers.hhs.gov/ohrp/categories/1569. Accessed August 24, 2011.

[20] Can I analyze data that are not individually identifiable, such as medication databases stripped of individual patient identifiers, for research purposes without having to apply the HHS protection of human subjects regulations? "Quality Improvement Activities – FAQs." Office for Human Research Protections. U.S. Department of Health and Human Services. http://answers.hhs.gov/ohrp/categories/1569. Accessed August 24, 2011.

[21] Casarett D, Karlawish JH, Sugarman J. Determining when quality improvement initiatives should be considered research: proposed criteria and potential implications. JAMA. 2000 May 3;283(17):2275-80.

[22] Dokholyan RS, Muhlbaier LH, Falletta JM, Jacobs JP, Shahian D, Haan CK, Peterson ED. Regulatory and ethical considerations for linking clinical and administrative databases. Am Heart J. 2009 Jun;157(6):971-82.

[23] Casarett D, Karlawish JH, Sugarman J. Determining when quality improvement initiatives should be considered research: proposed criteria and potential implications. JAMA. 2000 May 3;283(17):2275-80.

[24] Ezekiel J. Emanuel, M.D., Ph.D., and Jerry Menikoff, M.D., J.D. Emanuel EJ, Menikoff J. Reforming the regulations governing research with human subjects. N Engl J Med 2011; 365:1145-1150.

[25] Hannan EL, Kilburn H, Jr., O'Donnell JF, Lukacik G, Shields EP. Adult open heart surgery in New York State. An analysis of risk factors and hospital mortality rates. JAMA 1990 Dec 5;264(21):2768-74.

[26] Hannan EL, Kumar D, Racz M, Siu AL, Chassin MR. New York State's Cardiac Surgery Reporting System: four years later. Ann Thorac Surg 1994 Dec;58(6):1852-7.

[27] O'Connor GT, Plume SK, Olmstead EM, Coffin LH, Morton JR, Maloney CT, et al. A regional prospective study of in-hospital mortality associated with coronary artery bypass grafting. The Northern New England Cardiovascular Disease Study Group. JAMA 1991 Aug 14;266(6):803-9.

[28] O'Connor GT, Plume SK, Olmstead EM, Coffin LH, Morton JR, Maloney CT, et al. Multivariate prediction of in-hospital mortality associated with coronary artery bypass graft surgery. Northern New England Cardiovascular Disease Study Group. Circulation 1992 Jun;85(6):2110-8.

[29] Grover FL, Hammermeister KE, Shroyer AL. Quality initiatives and the power of the database: what they are and how they run. Ann Thorac Surg 1995 Nov;60(5):1514-21.

[30] Grover FL, Johnson RR, Marshall G, Hammermeister KE. Factors predictive of operative mortality among coronary artery bypass subsets. Ann Thorac Surg 1993 Dec;56(6):1296-306.

[31] Grover FL, Johnson RR, Shroyer AL, Marshall G, Hammermeister KE. The Veterans Affairs Continuous Improvement in Cardiac Surgery Study. Ann Thorac Surg 1994 Dec;58(6):1845-51.

[32] Edwards FH, Grover FL, Shroyer AL, Schwartz M, Bero J. The Society of Thoracic Surgeons National Cardiac Surgery Database: current risk assessment. Ann Thorac Surg 1997 Mar;63(3):903-8.

[33] Edwards FH, Clark RE, Schwartz M. Coronary artery bypass grafting: the Society of Thoracic Surgeons National Database experience. Ann Thorac Surg 1994 Jan;57(1):12-9.

[34] Grover FL, Shroyer AL, Hammermeister K, Edwards FH, Ferguson TB, Jr., Dziuban SW, Jr., et al. A decade's experience with quality improvement in cardiac surgery using the Veterans Affairs and Society of Thoracic Surgeons national databases. Ann Surg 2001 Oct;234(4):464-72.

[35] Blumberg MS. Comments on HCFA hospital death rate statistical outliers. Health Care Financing Administration. Health Serv Res 1987 Feb;21(6):715-39.

[36] O'Connor GT, Plume SK, Olmstead EM, Morton JR, Maloney CT, Nugent WC, et al. A regional intervention to improve the hospital mortality associated with coronary artery bypass graft surgery. The Northern New England Cardiovascular Disease Study Group. JAMA 1996 Mar 20;275(11):841-6.

[37] Ferguson TB, Jr., Peterson ED, Coombs LP, Eiken MC, Carey ML, Grover FL, et al. Use of continuous quality improvement to increase use of process measures in patients undergoing coronary artery bypass graft surgery: a randomized controlled trial. JAMA 2003 Jul 2;290(1):49-56.
[38] Tu JV, Donovan LR, Lee DS, Wang JT, Austin PC, Alter DA, et al. Effectiveness of public report cards for improving the quality of cardiac care: the EFFECT study: a randomized trial. JAMA 2009 Dec 2;302(21):2330-7.
[39] Hibbard JH, Stockard J, Tusler M. Does publicizing hospital performance stimulate quality improvement efforts? Health Aff (Millwood) 2003 Mar;22(2):84-94.
[40] Fung CH, Lim YW, Mattke S, Damberg C, Shekelle PG. Systematic review: the evidence that publishing patient care performance data improves quality of care. Ann Intern Med 2008 Jan 15;148(2):111-23.
[41] Hannan EL, Kumar D, Racz M, Siu AL, Chassin MR. New York State's Cardiac Surgery Reporting System: four years later. Ann Thorac Surg 1994 Dec;58(6):1852-7.
[42] Hannan EL, Kilburn H, Jr., Racz M, Shields E, Chassin MR. Improving the outcomes of coronary artery bypass surgery in New York State. JAMA 1994 Mar 9;271(10):761-6.
[43] Hannan EL, Siu AL, Kumar D, Kilburn H, Jr., Chassin MR. The decline in coronary artery bypass graft surgery mortality in New York State. The role of surgeon volume. JAMA 1995 Jan 18;273(3):209-13.
[44] Ghali WA, Ash AS, Hall RE, Moskowitz MA. Statewide quality improvement initiatives and mortality after cardiac surgery. JAMA 1997 Feb 5;277(5):379-82.
[45] Peterson ED, Delong ER, Jollis JG, Muhlbaier LH, Mark DB. The effects of New York's bypass surgery provider profiling on access to care and patient outcomes in the elderly. J Am Coll Cardiol 1998 Oct;32(4):993-9.
[46] Guru V, Fremes SE, Naylor CD, Austin PC, Shrive FM, Ghali WA, et al. Public versus private institutional performance reporting: what is mandatory for quality improvement? Am Heart J 2006 Sep;152(3):573-8.
[47] Hannan EL, Sarrazin MS, Doran DR, Rosenthal GE. Provider profiling and quality improvement efforts in coronary artery bypass graft surgery: the effect on short-term mortality among Medicare beneficiaries. Med Care 2003 Oct;41(10):1164-72.
[48] Vladeck BC, Goodwin EJ, Myers LP, Sinisi M. Consumers and hospital use: the HCFA "death list". Health Aff (Millwood) 1988;7(1):122-5.
[49] Hannan EL, Stone CC, Biddle TL, DeBuono BA. Public release of cardiac surgery outcomes data in New York: what do New York state cardiologists think of it? Am Heart J 1997 Jul;134(1):55-61.
[50] Chassin MR. Achieving and sustaining improved quality: lessons from New York State and cardiac surgery. Health Aff (Millwood) 2002 Jul;21(4):40-51.
[51] Schneider EC, Epstein AM. Use of public performance reports: a survey of patients undergoing cardiac surgery. JAMA 1998 May 27;279(20):1638-42.
[52] Schneider EC, Epstein AM. Influence of cardiac-surgery performance reports on referral practices and access to care -- a survey of cardiovascular specialists. N Engl J Med 1996 Jul 25;335(4):251-6.
[53] Mukamel DB, Weimer DL, Mushlin AI. Interpreting market share changes as evidence for effectiveness of quality report cards. Med Care 2007 Dec;45(12):1227-32.
[54] Mukamel DB, Mushlin AI. The impact of quality report cards on choice of physicians, hospitals, and HMOs: a midcourse evaluation. Jt Comm J Qual Improv 2001 Jan;27(1):20-7.
[55] Romano PS, Zhou H. Do well-publicized risk-adjusted outcomes reports affect hospital volume? Med Care 2004 Apr;42(4):367-77.
[56] Hibbard JH, Peters E. Supporting informed consumer health care decisions: data presentation approaches that facilitate the use of information in choice. Annu Rev Public Health 2003;24:413-33.
[57] Hibbard JH, Peters E, Slovic P, Finucane ML, Tusler M. Making health care quality reports easier to use. Jt Comm J Qual Improv 2001 Nov;27(11):591-604.
[58] Kaiser Family Foundation 2008 Update on Consumers' Views of Patient Safety and Quality Information. 2008. Henry J. Kaiser Family Foundation.

http://search.kff.org/gsaresults/search?site=KFForgnopdfs&filter=0&output=xml_no_dtd&client=kff&sp=kff&getfields=*&q=7819&no_pdf=1. Accessed September 16, 2011.
[59] Shahian DM, Normand SL, Torchiana DF, Lewis SM, Pastore JO, Kuntz RE, et al. Cardiac surgery report cards: comprehensive review and statistical critique. Ann Thorac Surg 2001 Dec;72(6):2155-68.
[60] Green J, Wintfeld N. Report cards on cardiac surgeons. Assessing New York State's approach. N Engl J Med 1995 May 4;332(18):1229-32.
[61] Jones RH. In search of the optimal surgical mortality. Circulation 1989 Jun;79(6 Pt 2):I132-I136.
[62] Lee TH, Torchiana DF, Lock JE. Is zero the ideal death rate? N Engl J Med 2007 Jul 12;357(2):111-3.
[63] Dranove D, Kessler D, McClellan M, Satterthwaite M. Is more information better? The effects of "report cards" on health care providers. Journal of Political Economy 2003;111:555-88.
[64] Omoigui NA, Miller DP, Brown KJ, Annan K, Cosgrove D, III, Lytle B, et al. Outmigration for coronary bypass surgery in an era of public dissemination of clinical outcomes. Circulation 1996 Jan 1;93(1):27-33.
[65] Burack JH, Impellizzeri P, Homel P, Cunningham JN, Jr. Public reporting of surgical mortality: a survey of New York State cardiothoracic surgeons. Ann Thorac Surg 1999 Oct;68(4):1195-200.
[66] Schneider EC, Epstein AM. Influence of cardiac-surgery performance reports on referral practices and access to care -- a survey of cardiovascular specialists. N Engl J Med 1996 Jul 25;335(4):251-6.
[67] Li Z, Carlisle DM, Marcin JP, Castellanos LR, Romano PS, Young JN, et al. Impact of Public Reporting on Access to Coronary Artery Bypass Surgery: The California Outcomes Reporting Program. Ann Thorac Surg 2010 Apr 1;89(4):1131-8.
[68] Bridgewater B, Grayson AD, Brooks N, Grotte G, Fabri BM, Au J, et al. Has the publication of cardiac surgery outcome data been associated with changes in practice in northwest England: an analysis of 25,730 patients undergoing CABG surgery under 30 surgeons over eight years. Heart 2007 Jun;93(6):744-8.
[69] Moscucci M, Eagle KA, Share D, Smith D, De Franco AC, O'Donnell M, et al. Public reporting and case selection for percutaneous coronary interventions: an analysis from two large multicenter percutaneous coronary intervention databases. J Am Coll Cardiol 2005 Jun 7;45(11):1759-65.
[70] Apolito RA, Greenberg MA, Menegus MA, Lowe AM, Sleeper LA, Goldberger MH, et al. Impact of the New York State Cardiac Surgery and Percutaneous Coronary Intervention Reporting System on the management of patients with acute myocardial infarction complicated by cardiogenic shock. Am Heart J 2008 Feb;155(2):267-73.
[71] Resnic FS, Welt FG. The public health hazards of risk avoidance associated with public reporting of risk-adjusted outcomes in coronary intervention. J Am Coll Cardiol 2009 Mar 10;53(10):825-30.
[72] Clarke S, Oakley J. Informed consent and surgeons' performance. J Med Philos 2004 Feb;29(1):11-35.
[73] Clarke S, Oakley J. Informed consent and clinician accountability: the ethics of report cards on surgeon performance. Cambridge: Cambridge University Press; 2007.
[74] Drozda JP, Jr., Hagan EP, Mirro MJ, Peterson ED, Wright JS. ACCF 2008 health policy statement on principles for public reporting of physician performance data: A Report of the American College of Cardiology Foundation Writing Committee to develop principles for public reporting of physician performance data. J Am Coll Cardiol 2008 May 20;51(20):1993-2001.
[75] Krumholz HM, Brindis RG, Brush JE, Cohen DJ, Epstein AJ, Furie K, et al. Standards for statistical models used for public reporting of health outcomes: an American Heart Association Scientific Statement from the Quality of Care and Outcomes Research Interdisciplinary Writing Group: cosponsored by the Council on Epidemiology and Prevention and the Stroke Council. Endorsed by the American College of Cardiology Foundation. Circulation 2006 Jan 24;113(3):456-62.
[76] NQF Evidence Task Force Final Report. http://www.qualityforum.org/Measuring_Performance/Improving_NQF_Process/Evidence_Task_Force.aspx. Accessed April 23, 2011.

[77] NQF measure selection. http://www.qualityforum.org/Measuring_Performance/ABCs/What_NQF_Endorsement_Means.aspx. Accessed July 22, 2011.
[78] NQF measure testing. http://www.qualityforum.org/Measuring_Performance/Improving_NQF_Process/Measure_Testing_Task_Force_Final_Report.aspx. Accessed July 23, 2011.
[79] Dimick JB, Welch HG, Birkmeyer JD. Surgical mortality as an indicator of hospital quality: the problem with small sample size. JAMA 2004 Aug 18;292(7):847-51.
[80] Hofer TP, Hayward RA, Greenfield S, Wagner EH, Kaplan SH, Manning WG. The unreliability of individual physician "report cards" for assessing the costs and quality of care of a chronic disease. JAMA 1999 Jun 9;281(22):2098-105.
[81] Fung V, Schmittdiel JA, Fireman B, Meer A, Thomas S, Smider N, et al. Meaningful variation in performance: a systematic literature review. Med Care 2010 Feb;48(2):140-8.
[82] Li Q, Glynn RJ, Dreyer NA, Liu J, Mogun H, Setoguchi S. Validity of claims-based definitions of left ventricular systolic dysfunction in Medicare patients. Pharmacoepidemiol Drug Saf. 2011 Jul;20(7):700-8.
[83] Porter ME. What is value in health care? N Engl J Med 2010 Dec 23;363(26):2477-81.
[84] Jacobs JP, Edwards FH, Shahian DM, Prager RL, Wright CD, Puskas JD, et al. Successful linking of the society of thoracic surgeons database to social security data to examine survival after cardiac operations. Ann Thorac Surg 2011 Jul;92(1):32-9.
[85] Jacobs JP, Edwards FH, Shahian DM, Haan CK, Puskas JD, Morales DL, et al. Successful linking of the Society of Thoracic Surgeons adult cardiac surgery database to Centers for Medicare and Medicaid Services Medicare data. Ann Thorac Surg 2010 Oct;90(4):1150-6.
[86] Bufalino VJ, Masoudi FA, Stranne SK, Horton K, Albert NM, Beam C, et al. The American Heart Association's recommendations for expanding the applications of existing and future clinical registries: a policy statement from the American Heart Association. Circulation 2011 May 17;123(19):2167-79.