

Quality of Care Issues

The Definition of Quality Care:

Much of the quality-of-health-care literature is focused on a narrow definition of quality as the reduction of medical errors. It may also focus on the reduction of variance in treatment or outcomes. This narrow view of quality is sometimes called “patient safety”. This definition is most consistent with the quality efforts in manufacturing (TQM, Baldrige, etc.). It is also consistent with a common view that the big issues in healthcare are quality, access, and price. In other words, quality and access are different issues. Other articles are much broader and include such things as access to healthcare as part of the definition of quality. This definition is reasonable if one is trying to explain how US healthcare ranks outside the top ten countries in the world as measured by common quality indicators such as life expectancy and infant mortality. The disadvantage of including issues like access and financing in quality discussions is that quality loses meaning. Any article about healthcare becomes an article about quality. In general, we tried to stay with the more narrowly focused articles.

Many of these articles refer to “industry” or “industrial” quality efforts. By that, they mean quality efforts outside the healthcare industry. The error and variance reduction in, for instance, manufacturing and aviation, is a record to be emulated in healthcare. Many articles refer to “industrial-strength quality”. Most such references here are in the Patient Safety subsection. A common theme of these articles is to determine the “barriers to quality” that have kept healthcare from achieving the quality improvements of other industries. Barriers commonly mentioned include perverse financial incentives, reporting problems, education problems, and a lack of leadership. We find many articles addressing these specifics and group them below.

Another definitional typology of quality articles is the Donabedian approach dating to the sixties. He said quality should be measured in structure, process, and outcome. Structure studies look at such measures as the nurse-to-patient ratio, the educational level of the nurses, the degree of specialized education of doctors, whether the hospital has a quality officer or an intensive care specialist, etc. Some of the articles in the Leadership Section and some of the articles on nursing or intensive care units fit this category. Process studies look at whether guidelines or best practices are followed. All of the studies in the Standards and Education Section look at process as well as many of the HMO studies focused on inoculations or patient access. Outcome studies are most directly concerned with the patient and presumably the most important studies. However, the structure and process studies are more actionable. They not only reveal shortcomings but they tell the provider or plan what to work on. Deficient outcomes often leave a provider with no idea what to do next (except to study structure and process). Also, outcomes may take years to materialize. (The WHI study of hormone replacement was halted after five years because heart disease mortality was increased rather than decreased. Still, they could not refute other studies or speculation that the long-term impact could have become positive.) Outcome studies in the following articles concern risk-adjusted mortality results, complication rates, readmission rates, expense, quality of life, or patient satisfaction. Quality of life studies include functional and cognitive results. Some studies (like

mammography accuracy or inoculation rate) fall between process and outcome. The risk adjustment section of this bibliography also has numerous articles on quality outcome measurement.

To increase thoroughness, four of these articles are literature reviews. One is closer to a meta analysis.

The Institute of Medicine's (IOM) Reports

The 2000 publication of the IOM's "To Err is Human: Building a Safer Health System" report focused on patient safety and was of great significance. It seems to have synthesized the best of earlier quality efforts and set a standard to which all quality articles since must refer. It defined errors as overuse of inefficient care, underuse of efficient care, and misuse of care. It received much media and academic attention. To take advantage of its success, the IOM quickly (2001) released "Crossing the Quality Chasm: A New Health System for the 21st Century". The latter uses a broader definition of quality and includes considerable opinion on financing and access. Many of the articles abstracted here refer to "the IOM reports" or to specifically one of these reports. They (or at least the earlier report) are often credited with getting quality back on the public agenda after a ten-year hiatus. We abstract their executive summaries in Kohn and Berwick. The reports may be ordered as books or read online at www.nap.edu.

Measurement and Reporting

Beyond defining quality, we must measure it to improve it. Further, if we are to learn from our errors, they must be analyzed and reported to management. If all physicians are to learn from others' errors, reporting must be widespread. If market forces are to be enlisted to demand quality, then plans, employers, and even patients must be able to obtain useful information on provider quality.

Actuaries will find it interesting that risk-adjusted mortality is the most popular quality measure. (Patient satisfaction is perhaps more popular but not so specifically quality-oriented.) The tool is not without its critics, nor is it new. Krauss quotes Florence Nightingale on the need for risk adjustment and on the idea that comparing mortality rates would be more appropriate if the purpose of hospitals was to kill patients. Chassin discusses a very successful program (as is also noted in several of these articles) in which the New York Department of Health began reporting cardiac surgery mortality rates by hospital. The worst hospitals quickly improved or quit doing cardiac surgery. The specific intent was to concentrate the surgeries in high volume providers that most research agreed had safer care. Three articles describe the growing amount of provider-specific quality information available on the Internet. Most research agrees that patients are not yet using this information to make provider decisions and some websites state that provider decision-making is not their purpose. Bates reviews what is available and describes disease-specific bulletin boards as a more efficient extension of word-of-mouth, the traditional patient information source. Krumholz is concerned with the quality of information on the Internet (which is usually proprietary in data and method) and does a test of the most popular website, Healthgrades.com. Like most risk-adjustment methods, the Healthgrades rankings are found to distinguish between large groups but not

individual providers. Krumholz does not think Healthgrades information can serve its intended purpose of helping patients choose providers. Scalise is similar to Krumholz in expressing great concern for Internet sources. She provides a list of such websites and their characteristics. Lansky calls for more public disclosure of provider quality. He says the government should require it in its role as a purchaser of healthcare. He calls the State Children's Health Insurance Program's required reporting a success and a model for any future government purchases such as Medicare prescription drugs. Rosner is a disturbing piece. It says physicians have an obligation to reveal their errors to their patients and cites standards in their codes. Then Rosner describes situations in which it is in the best interest of the patient if errors are not revealed. For instance, a lawsuit would disturb the necessary continuity of care, and revelation of errors causes patients in general not to trust their physicians. These cultural values are one reason errors persist according to quality advocates. It demonstrates the basis of their arguments for revised roles for physicians as team players in a system. Today they are the ones responsible for all knowledge and all success or failure.

We finish this measurement and reporting section with an article on how reported information affects the buyers of health care. This will lead us to the section on financing healthcare. Hargraves says that employers haven't used quality data when choosing health plans for their employees because the available information is not useful. They prefer provider information to plan information because of plan overlap. He discusses the demise of employer coalitions and the growing attempts to make employees better purchasers of health care.

Financing

Born studies the relationship of HMO profitability and quality. He says the economists (who say performance leads to profits) win over the policy analysts (who express concern that profits and quality are in conflict.). He cautions that the poor financial results achieved by many HMOs is a warning sign for future quality. Cox studies the satisfaction of different populations with Medicare Plus Choice plans. The disabled and frail elderly are concerned with access. Minority status and education level did not predict satisfaction. Davis reviews the literature on how HMOs affect quality. Most articles find them too focused on cost cutting. Davis calls for government standards. Landon describes an extensive study of Medicare Plus Choice beneficiaries. He determines how plan features affect their satisfaction. Miller provides something beyond a literature review but less than a meta analysis. Beginning with many more articles on HMO performance, he narrows it down to 79 that clearly test the relative performance of HMOs and other plans on some criteria. He counts the number of times the HMO result is more positive. HMOs use fewer resources but have access problems and lower patient satisfaction. Shi studies how the type of insurance coverage affects primary care. His significant results are that FFS patients receive the best care and the uninsured the worst. HMO patients do receive the most preventative care but otherwise fall behind FFS patients on a number of issues. Dembe summarizes new research on the quality of health care financed by workers' compensation programs. Injured employees usually have less choice of provider and less satisfaction with their care, but the data is very limited. Allesandrinni studies FFS and

HMO impacts on childhood immunizations in a small Medicaid setting and finds no difference.

Patient Safety

Ayanian compares coronary heart disease (CHD) treatment quality in the US and UK. CHD is a leading cause of death in each country but has higher death rates in the UK. As might be expected, the US has a mix of high quality care and poor or no care by population group. The UK provides inadequate resources resulting in waiting lines. Each country has guidelines for care that are not followed. Information, less fragmentation, and better incentives are keys to improvement. The UK can have quicker success with top-down control. Beecher (Health Affairs 20-3) describes a systems approach to quality improvement. Physicians must be retrained to function on a team. Systems and processes must be redesigned to anticipate inevitable human errors and prevent them. Incentives must be found to encourage quality since the market has not done it. Fee for service rewards overuse errors, capitation rewards underuse errors, and misuse errors increase hospital and physician income. Coye discusses the IOM reports and many other quality initiatives. She thinks perverse incentives, lack of information, lack of leadership, and lack of demand have allowed poor quality to persist but she sees hope in baby boomers (the activist generation) and in the slow spread of serious quality efforts. Shannon studies perceptions of quality by the physician, nurse, and patient. Their differences and the implications are discussed. Goldstein studies the applicability of the Baldrige Award criteria to healthcare. Some but not all Baldrige criteria are determined to relate to quality outcomes. Langemo studied the effect of processes of nursing care on patient and nurse satisfaction. Newhouse wrote the most economically oriented and pessimistic article in this section. He thinks the lack of quality is an inefficiency problem (failure to produce the highest quality output for a given input). He sees inefficiency as the natural result of the economics. There are great amounts of uncertainty, information asymmetry, and moral hazard. There are perverse financial incentives, rapid technological change, and extensive government involvement. All these cause inefficiency which, in medical care, manifests itself as poor quality. Glance tests different risk-adjustment methods for ICU mortality to see how the methods affect the quality rankings. Fortunately, there is reasonable agreement. Devers studied quality improvement efforts at a number of hospitals. She largely confirmed the IOM reports' discussion of barriers to quality. The Women's Health Initiative (WHI) reports on a major study of the impact of hormone replacement therapy and concludes that its net impact is detrimental to postmenopausal women. Douglas studied the use and misuse of long-term ventilation. Her article is noteworthy for the many and carefully documented outcomes measures but she begins, like most, with mortality.

Standards and Education

Many articles say there is a need for more research to establish clinical guidelines and standards. However, others point out that the existing guidelines are not followed very well. Borbas studies guideline diffusion and performs an intervention to spread the use of medical guidelines in Minnesota. He specifically studies the role of opinion leaders in the process. Chang studies the difference between nursing guidelines and actual practice. At

least twenty percent of care was out of guideline in each location studied. Smaller hospitals and those serving poorer neighborhoods performed most poorly. Ferris performs a large literature review. He seeks studies using randomized control trials to test a quality improvement intervention for pediatric care. The number of such studies is increasing rapidly, showing researcher and funding interest. Jencks describes a very large CMS backed study of compliance with guidelines in Medicare fee-for-service settings. State-by-state information is given as well as the overall poor result of 69% compliance. Kiefe, in another CMS-backed study, describes a randomized control trial of a quality intervention to improve guideline compliance. Of interest to actuaries, the latter two discuss, and even attempt to estimate, the lives that could be saved with better guideline compliance. Esserman says that mammography readers in the UK are much more accurate than in the US due to their much higher volume. She argues for a more centralized system in the US which would allow greater specialization.

Leadership

Many articles that list the causes of poor quality mention a lack of leadership. There seems to be a lack of incentive for anyone to step forward. Providers and physicians that greatly reduce errors will reduce their income. Beecher (Health Affairs 20-5) discusses errors in medicine and this need for leadership. Competition led other industries to improve quality, but there seems to be no competition for US health care. The best hope may be consumer movements. McGlynn describes how much more dangerous medical care is than flying or riding on Firestones but the public's concern is with the latter. She calls for government action and for those funding research to specify treatments that can be and will be delivered accurately. We end with a good article for a first reading in health care quality. Sprague provides an overview of the efforts of various groups to advance health care quality. She gives a good summary of the many government initiatives as well as public-private cooperation.

Alessandrini, E. A., K. N. Shaw, et al. (2001). "Effects of Medicaid Managed Care on Quality: Childhood Immunizations." Pediatrics **107**: 1335-1342.

Keywords: Immunizations, Medicaid, provider funding

Purpose: The authors studied the impact of insurance plan (fee-for-service-FFS v. managed care-MC) on childhood immunizations in a Medicaid population.

Data: From over 644 births, 76 FFS and 437 MC were tracked for two years and their immunization status was determined.

Methods: All births were paid for by Medicaid in one large Philadelphia hospital. The MC cases came from three counties in which Medicaid had contracted with MC plans and the FFS cases came from three other counties. Multiple logistic regression was used to determine factors predictive of immunization status.

Results: 40% of US births are paid for by Medicaid. 25% of US children are enrolled in Medicaid. The overall quality of care provided children has been shown to be well-predicted by the children immunization rate. The key result in this study is that the MC and FFS babies had almost identical immunization results (73% MC and 72.4% FFS). Factors that increased the odds of immunization were the firstborn child and adequate prenatal care. Factors that decreased the odds of immunization were a father living in the home and private office based primary care. The immunizations are free under the federal Vaccine for Children program. The doctor need only tell the Department of Public Health how many patients he has under Medicaid.

Uses: The supposed advantage of the MC plan in quality of preventative care was not evident in this Medicaid setting.

Limitations: Only one hospital in one city was studied although several MC plans were involved. Local differences could have overwhelmed MC effects.

Ayanian, J. Z. and T. J. Quinn (2001). "Quality Of Care For Coronary Heart Disease In Two Countries: The United States and England each have implemented strategies that might prove useful to the other." Health Affairs **20**(3).

Keywords: CHD, fragmentation, information, quality

Purpose: The authors compare efforts to improve the quality of care of coronary heart disease (CHD) in England and the US. CHD is responsible for 20% of deaths in the US but the CHD mortality rate is even higher (by 36%) in England.

Data/Methods: Much secondary data is provided but not original research.

Results: In the US, there is more than an adequate supply of care available for insured patients. In England, the supply is limited, requiring waiting lines for CHD treatment. Some receive less than optimal care. The US spends 10% of its healthcare dollars on CHD but quantity rather than quality is emphasized. Care is fragmented. Subpopulations are underserved. Clinical guidelines for care have been used in each country. Compliance is usually voluntary and there is little evidence that they have changed patterns of care. In the US, data is available on the CHD mortality results of physicians and hospitals. In England, public performance reporting is more limited. This information provides benchmarking for medical facilities as well as selection criteria for patients. The market may respond to quality if information is available. Evidence shows that employers use quality information more than managed care plans do when making contracting

decisions. The authors think the future of quality care will depend on the development of less fragmented systems using improved information, improved organizational culture, and appropriate incentives. Risk-adjusted quality information is a necessity. England can use more top-down approaches to improving care while the US government agencies can only provide feedback and guidelines. More research is needed on the impact of incentives for quality.

Uses: This provokes thought on how improvement can be made at the macro level.

Limitations: This is an opinion piece based on considerable secondary data.

Bates, D. W. and A. A. Gawande (2000). "The Impact of the Internet on Quality Measurement; Word-of-mouth advice about providers is gaining respectability through the web." Health Affairs **19**(6).

Keywords: Internet, provider information, quality

Purpose: The authors analyze the impact of the Internet on the availability of quality information to patients.

Data/Methods: Internet information on health care is reviewed. There are 17,000 websites devoted to healthcare.

Results: The authors view the Internet as an extension and improvement on the traditional source of quality information: word of mouth. The inherent weaknesses are being slowly overcome by the movement to report cards and physician profiling. The authors discuss a number of online and other information services. They argue that risk-adjusted outcomes comparison is not enough. Resource availability and process are also important. The information is adequately timely and few patients use it. However, there is some evidence that providers are aware of their performance data and they strive to improve it. Surprisingly, Bates thinks the online chat rooms, bulletin boards, and list serves are more often used for advice on providers than the information service providers. This is an extension of older word-of-mouth methods. And there is some evidence that reputation does distinguish quality. Data on medical errors is generally still unavailable. Disciplinary actions against physicians are unavailable. Some information on the Internet is biased. Legal liability for misinformation on the Internet is unclear.

Uses: This is descriptive and does not advocate positions.

Limitations: The authors cite a number of problems with US healthcare that even an improved Internet will not address: access for the poor, access to a choice of plans and providers,

Becher, E. C. and M. R. Chassin (2001). "Improving the Quality of Healthcare: Who Will Lead? To date, no party has stepped into the leadership role that the quality movement so badly needs." Health Affairs **20**(5).

Keywords: Incentives, leadership, quality

Purpose: The authors make the case that the disappointing lack of progress on quality issues in healthcare is due to lack of leadership.

Data/Methods: This is an opinion piece.

Results: The authors describe the lack of quality in healthcare. They document the extent of overuse, underuse, and misuse of medical care. They review the history

of concern for poor quality, peaking in 1988 and again in 1999, with little progress made. They describe how none of the participants in healthcare has a financial incentive to attack poor quality in each of its three forms. Physicians and hospitals don't want underuse especially if getting fee for service. Employers, HMOs, and insurers don't want overuse. Providers do not want misuse revealed or reported. Providers still tend to blame the one who made the mistake rather than the system. The authors describe how competition has led to major quality improvement in other industries, but they don't see a source of competition for US healthcare. They don't see government bringing about quality change because it could not even successfully produce guidelines. Academic medicine shows little interest. Thus they look for a leader. They describe environmental and automotive safety consumer movements that they say changed industries, and they look for a similar leader to emerge in US healthcare.

Uses: This could inspire the insurance industry (which the authors think is focused on cost reduction) or actuaries to attempt this leadership role.

Limitations: This is an opinion piece with little secondary data.

Becher, E. C. and M. R. Chassin (2001). "Improving Quality, Minimizing Error: Making It Happen; A five point plan and why we need it." Health Affairs **20**(3).

Keywords: Medical errors, quality improvement

Purpose: The authors propose a program to overcome the structural barriers to improvement in US healthcare quality.

Data/Methods: This is an opinion piece.

Results: The authors say that medical errors affect millions of patients each year. They review theories of human error and especially industrial error. The latter are errors arising in complex interaction of many people. In complex situations errors increase. In complex situations involving many participants errors become inevitable. Systems must be designed to anticipate, detect, and correct errors in process. Medical training and historical roles are at odds with this process or systems improvement thinking. Doctors, in particular, have not been trained to work in a team with other doctors or other professionals. Existing organization, processes, and systems do not facilitate communication. Status quo thinking blames and punishes an individual for each error. The incentive is to hide errors. There is no financial incentive to improve quality. Providers would have to make the investment to improve quality, and third party payers would reap most of the savings. Neither patients nor employers are demanding quality. There is little data on real error rates but what there is indicates that errors are increasing. This might be expected as medical care becomes more complex and there is no organizational improvement. They propose five directions for investment in quality: 1. The public must be educated about the current lack of quality so as to increase demand for quality. 2. Government invests heavily in medicine. Investments in process improvement could be more effective than further investments in more complex procedures or equipment. 3. The payment systems must be revised to reward quality. Fee for service rewards overuse errors and capitation rewards underuse errors. There is no current system rewarding quality improvement. 4. State licensure should target the few doctors who repeatedly make serious errors. 5. The leadership void in quality improvement must be filled.

Uses: Advocates for quality improvement will find some possible direction here.

Limitations: Little data is presented but that will always be the case so long as error reporting is suppressed.

Berwick, D. M. (2002). "A User's Manual For The IOM's 'Quality Chasm' Report." Health Affairs **21**(1): 80-90.

Keywords: Medical errors, quality improvement, quality measurement

Purpose: The express purpose is to do the executive summary that the Crossing the Quality Chasm did not contain.

Data/Methods: Not applicable.

Results: Berwick was a lead author of the Crossing the Quality Chasm report so he is well-qualified to write this executive summary. He says the underlying framework of the initial study was not described in the book and that makes the book a difficult read. This summary reveals the framework. He also makes many references to the IOM's earlier effort on patient safety, To Err is Human. That earlier report received widespread promotion and it is clear that the IOM wanted to get more of their ideas into the discussion that the To Err is Human report started. The missing framework is that the most important changes will occur at Level A, the realm the patient experiences. Level B is the subsystem that directly deals with the patient. If the patient has surgery, it is the surgeons, nurses, and other specialists involved as well as their facilities, equipment, and processes. Level C is the local institution housing the subsystem. Level D is the environment for the local institution. It consists of laws, regulations, payment systems, etc. Level A is "True North" and all other levels have their quality judged by the impact at level A. They call for six Aims for Improvement: 1. Safety. 2. Effectiveness, meaning the avoidance of both overuse of ineffective care and underuse of effective care. 3. Care should be centered around the patient and respect the patient's wishes. 4. Timeliness. 5. Efficiency defined as avoiding waste of resources, ideas, and spirit. 6. Equity.

Uses: This is ammunition for those calling for major overhaul of US healthcare. Berwick says the current system is incapable of providing quality health care. Referring to the To Err is Human report, "current rates of injury from care are inherent properties of current system design and that safer care will require new designs". Redesign suggestions are from systems theory and especially, complex adaptive systems theory.

Limitations: This is the executive summary version.

Borbas, C., N. Morris, et al. (2000). "The Role of Clinical Opinion Leaders in Guideline Implementation and Quality Improvement [Translating Guidelines Into Practice: Implementation and Physician Behavior Change]." Chest **118**(2 Supplement): 24s-32s.

Keywords: Guideline implementation, opinion leaders, quality improvement

Purpose: There is a significant lag between the publication of practice guidelines and their widespread adoption in clinics. The authors research the role of opinion leaders in this process.

Data: Practice guidelines for four drugs were presented to 37 Minnesota hospitals. Three were positive recommendations to use a new drug, and one was a

guideline to restrict the use of an older drug.

Methods: A randomized control trial was used. Patient records before and after intervention were reviewed. The intervention was identification and training of opinion leaders. Doctors were also surveyed about their knowledge of the guidelines.

Results: The authors review diffusion innovation research, adult learning theory, and social influence theory as frameworks. They acknowledge that concepts from sociology, psychology, communication, and marketing must inform education efforts. Specifically, they review the role of opinion leaders. In medicine, this role has been studied for 50 years. These leaders are local peers. They are not in authority. They are not innovators but are early adopters. Their great influence is the reason practice varies so widely in the US. Group norms, local values, and local practice realities determine acceptance of innovation. Leadership roles vary by system and disease. These leaders connect local practice to the new innovations. Six barriers to guideline adoption are identified and solutions are discussed.

Uses: This intended use is in the design of education programs about new practice guidelines. Educational materials are not enough. Sociological issues must be addressed. Local opinion leaders must be targeted.

Limitations: The authors state that it is easy to underestimate the complexity of the educational task.

Born, P. H. and C. J. Simon (2001). "Patients and Profits: The Relationship Between HMO Financial Performance and Quality of Care; Financial resources and profitability serve as an "early warning system" in monitoring the quality of care." Health Affairs **20** (2).

Keywords: HMO profitability, HMO quality

Purpose: The authors study the relationship of HMO profitability and quality. They refer to concern that these are in conflict. They look for other determinates of plan quality such as ownership status.

Data: About 200 HMOs are studied. Data is from 1997-8 National Committee for Quality Assurance (NCQA) Health Plan Employer Data Set (HEDIS). They acknowledge the debate over the usefulness of this data. They link it to financial information from the National Association of Insurance Commissioners (NAIC).

Methods: Multivariate analysis is used to determine the impact of the variables on quality. Numerous variables such as demographics and local market conditions are compared with financial performance.

Results: Ownership status does not predict quality. Past profitability is a positive indicator for later quality. This raises concern as many HMOs have records of past losses. Other significant variables show that nonwhite and urban demographics are associated with lower quality while competition and the presence of large employers is associated with higher quality.

Uses: The authors say they confirm the view of economists (profits lead to quality) rather than the view of many health policy analysts (profits and quality are in conflict). Financial losses could be an early warning sign of quality problems.

Limitations: Quality is measured by services provided rather than error rates.

Chang, B., J. Lee, et al. (2002). "Evaluating Quality of Nursing Care: The Gap Between

Theory and Practice." The Journal of Nursing Administration **32**(7/8): 405-418.

Keywords: Clinical practice variation, nursing, quality

Purpose: This article reviews variation in practice from theory and the degree of variation by type and locale of practice.

Data: Extensive case review is done by peer reviewers. Almost 600 cases are extensively reviewed but selected from a sample of 17,000 on which some data was compiled..

Methods: Regression on two illnesses showed determinates of quality nursing care.

Results: Previous studies indicate that the quality of nursing care has as great an impact on mortality as the quality of physician care. It has been studied much less. Less than adequate care was found in a large percentage of cases. It varied from under 20% for some diseases/functions to over 80% for others. The authors reviewed nursing education in the areas of poorer performance and found that they were not adequately emphasized. Diseases were limited to two: heart disease and cardiovascular accident, representing different types of care. Functions performed well were assessment and measuring vital signs. Functions performed poorly were planning treatments, following doctors plans, medication tracking, patient education and psychosocial management. Smaller, but not rural, hospitals had less adequate care. Hospitals serving high-poverty areas had poorer quality. Ownership was not a significant determinant.

Uses: The authors want this to be a baseline study so that updates will track quality improvement.

Limitations: The data is old, having been collected in the mid-1980s. The authors warn that since then, there has been a massive push to have lower-quality nurses do as much of this work as possible.

Chassin, M. R. (2002). "Achieving And Sustaining Improved Quality: Lessons From New York State And Cardiac Surgery; Deaths from cardiac surgery fell 41 percent over the first four years of New York's reporting program and continued to fall ever since." Health Affairs **21**(4).

Keywords: CABG, quality improvement, quality measures, volume

Purpose: Chassin describes New York's very successful efforts to improve the risk-adjusted mortality rate of coronary artery bypass graft (CABG) surgery.

Data/Methods: This is descriptive of the quality improvement program for CABG in NY. It is a case study.

Results: NY began measuring risk-adjusted CABG mortality by hospital and surgeon in 1989. They had tremendous improvement throughout the nineties. They quickly discovered that low volume hospitals and doctors had the worst results. Many programs were closed, and they have focused (through licensing) on developing a limited number of high volume CABG centers in major cities. Chassin reviews case studies of how improvement was accomplished at other hospitals. Generally, specialization of surgeons, other staff, facilities, and management were used. He says most of the improvement was in the worst hospitals. He is confident that neither consumers nor HMOs demanded this improvement. The incentive was from government or self imposed. Other factors are researchers to publish the results(!), government backing, and involvement of the state's leading

cardiac surgeons from the beginning.

Uses: This shows how the availability of good quality data is likely to lead to quality improvement even though much additional work, beyond the data, is required.

Limitations: Chassin doubts if the results are reproducible by another state. The strength of the NY Dept. of Health was key. Outcomes measurement does not usually show what needs to be done, just that something must be done. Critics say NY hospitals learned how to record illnesses to get high risk-adjusted expected mortality or they learned to send bad risks out of state. Chassin disputes both but his reasoning seems faulty.

Cox, D., K. Langwell, et al. (2001). "Assessing Medicare Health Plan Performance in Serving Beneficiary Subpopulations." Health Care Financing Review **22**(3): 85-99.

Keywords: HMO, patient satisfaction

Purpose: The authors review survey data on the degree of satisfaction HMO members express and how it varies for certain populations.

Data: 125,000 Medicare enrollees were surveyed in MMC-CHAPS. They came from 199 HMO plans.

Methods: Degree of satisfaction is regressed on independent variables such as plan type, beneficiary type, duration in the plan, and demographics.

Results: This is a popular research field. Many prior studies are reviewed. They are not identical to this study in that they do not single out these subgroups or they are targeted at providers rather than plans. Prior results are inconsistent. The disabled are particularly dissatisfied with plan access. The frail elderly are also dissatisfied. Minorities and lower educated varied little from whites and higher educated. The satisfaction of all groups tends to come together with duration in the plan. Prescription drug benefits did not improve satisfaction. The difference between fee-for-service plans and managed care plans was slight.

Uses: HCFA wants satisfaction measured and improved for Medicare Plus Choice plans and wants the needs of certain groups to be addressed. This is to fill that need and point in the direction of needed improvements.

Limitations: The authors express none but discuss several directions for further research.

Coye, M. (2001). "No Toyotas In Health Care: Why Medical Care Has Not Evolved To Meet Patients' Needs; Until payment policies reward quality improvement, providers will not place it at the core of their business strategy." Health Affairs **20**(6).

Keywords: Barriers to quality, quality improvement

Purpose: This article reviews many quality improvement initiatives in US healthcare, finding none of them adequate but seeing some signs of hope.

Data/Methods: This is not original research but literature review and opinion.

Results: The obstacles to improved quality in US healthcare are reviewed. Coye summarizes them by saying there has been no business case for quality improvement. She mentions many quality reports and efforts. Among these are the Institute of Medicine's (IOM) "To Err is Human" report, the Committee on Quality of Health Care in America's "Crossing the Quality Chasm" report, Health Plan Employer Data and Information Set (HEDIS) reports, the National

Committee for Quality Assurance requirements for HMOs, the employer-motivating Leapfrog initiative, the Institute for Health Care Improvement's "Pursuing Perfection" initiative, and the formation of the National Forum for Health Care Quality Measurement and Reporting. Her assessment of the barriers to quality are similar to other articles in this bibliography: perverse incentives, lack of information, lack of leadership, and lack of consumer demand.

Concerning the latter, she animatedly calls for a consumer advocate to inform the public that they are going for hospital rides in Corvairs, not Toyotas. She thinks the baby boomers, "the most activist consumer cohort in history," may demand change. Her most optimistic comments are reserved for small, individual efforts at quality improvement. She refers to innovators as only two percent of a profession. She calls early adopters the next 15% of a profession to change. Once the early adopters have changed, the rest of the profession follows. She says quality improvement desire and knowledge has spread well into the early adopter group.

Uses: She informs us of the current status of quality improvement efforts in US health care.

Limitations: She does not give detail on the smaller efforts. She says what has taken place is necessary but insufficient.

Davis, K. (1999). "Can the Market Ensure Quality Work without Government?" Journal of Health Politics, Policy and Law **24**(5).

Keywords: HMO quality, regulation, standards

Purpose: Davis reviews how HMOs have performed in increasing their quality and advocates a government role setting standards and providing information.

Data/Methods: This is an opinion piece.

Results: This focuses on quality in the HMOs' performance. Patient satisfaction and access to care are the key issues. Based on review of other published studies, plan performance varies widely. In theory, HMOs could drive provider quality but in practice they have not. They are too focused on cost cutting. Employers and employees have not used available information to select quality plans. Another problem is that most doctors are contracted with several HMOs so no one HMO can set standards for all of the patients. The government will increasingly set standards for reporting and performance because it is the largest purchaser of HMO services through Medicare and Medicaid.

Uses: The intention is to advocate a greater role for government in HMO regulation.

Limitations: No original data is provided.

Dembe, A. E., S. E. Fox, et al. (2002). "The RJWF Workers' Compensation Health Initiative: Findings And Strategies; As this initiative winds down, participants at a September 2001 meeting generate an agenda for promoting optimal care for work-related health conditions." Health Affairs **21**(1).

Keywords: Fragmentation, quality, Workers Compensation

Purpose: As the title notes, the findings from six million dollars of research funding by the Robert J. Woods Foundation are summarized. The meeting that summarized the findings and was the basis for this paper also set future directions for

workers' health care quality improvement and research.

Data/Methods: This is a general summary of findings and a collection of lists for quality improvement and research directions.

Results: Quality of care under workers' compensation laws is possibility even more challenging than under other financing. The fragmentation is still there and there is, if anything, greater pressure on cost control. There is the additional focus, which can be positive, of limiting disability and returning the sick or injured to productive status. Recent attempts at more coordinated care have shown the potential for twenty to thirty percent reductions in medical costs but at the expense of less patient satisfaction. Choice of provider is generally more limited under workers' compensation care than under other plans. This research has begun some efforts to collect data on workers' compensation nationally and to direct further research. The Agency for Healthcare Research and Quality (AHRQ) is encouraged to fund more research.

Uses: Many directions for further effort and research are provided.

Limitations: This is an overview of a fairly new field of research.

Devers, K. (2002). "From The Field; Quality Improvement By Providers: Market Developments Hinder progress; Patient-safety initiatives stand out as one area in which providers have taken steps to improve quality." Health Affairs **21**(5).

Keywords: Barriers to quality, hospital, quality improvement

Purpose: This is an update to earlier Community Tracking Study (CTS) results. Each two years, hospitals and clinical leaders are interviewed about quality initiatives.

Data: Forty-eight hospitals and forty-eight medical groups were studied, requiring 202 interviews. They are in 12 different cities.

Methods: Quality initiatives, quality results, and barriers to quality were probed.

Results: Between 1999 and 2001 quality programs were maintained but not increased. Pressure for quality came in the form of the two Institute of Medicine reports but market forces increased the barriers to quality. Quality efforts included specific staffing for quality, goal-setting, systems implementation, and guideline implementation. Devers refers to these activities as early stage quality improvement activities. Medical care is still far behind other industries in quality improvement. Barriers to quality improvement include increased financial pressure, a decline in tightly managed HMOs and capitation, strained relations between plans and providers, and a retrenchment of organized delivery systems. One favorable development was a small increase in financial rewards for quality. These were typically a 4-10% bonus for meeting HEDIS and satisfaction goals. Many plans were launching specific disease care management programs or adopting national services to perform this function. Providers were hostile to this as further fragmentation of the patient care system and communication. Patient safety programs showed some increase. Leadership and accountability, medication, and staffing were areas of improvement efforts. Barriers to these programs were cost, liability (fear of reporting errors), lack of research on such programs, and physician resistance. There is some indication that consumers became more aware of medical errors.

Uses: This confirms the IOM reports' conclusion that major changes are required.

Quality improvement has had periods of peaked interest in the past. Devers asks

if the current emphasis on patient safety can be sustained and can it lead to general quality improvement?

Limitations: The focus is large hospitals in metropolitan markets.

Douglas, S. L., B. J. Daly, et al. (2002). "Survival and quality of life: Short-term versus long-term ventilator patients." Critical Care Medicine **30**(12).

Keywords: Outcomes, quality of life, ventilation

Purpose: The authors describe and compare the outcomes of short and long-term ventilation (STV and LTV, respectively) intensive care unit (ICU) patients. They say there has been controversy about the benefits of LTV but definitions are inconsistent. Further, this is the first study that compares STV and LTV outcomes.

Data: 538 patients were followed one year after discharge. 47% died in the hospital and 65% died within one year.

Methods: The following outcomes were studied: mortality, quality of life (QOL), and expense for patients ventilated 24-96 hours (STV) and more than 96 hours (LTV). Quality of life was measured using the Sickness Impact Profile (SIP). Expenses were standardized at the Medicare reimbursement rates. Many covariates were studied. Patients were reexamined at two weeks, six months, and one year post discharge.

Results: Half of STV patients were discharged to their home (rather than to a nursing home or another hospital, etc) but only one fourth of the LTV discharges were to their home. STV patients who were discharged to a nursing home spent an average two months there before going home. LTV patients spent seven months there. Both LTV and STV patients had readmission rates of about 50%. The average expense for all care for STVs was \$70,000 and for LTVs \$131,000. The figures for only those who survived an year were \$92,000 for STVs and \$179,000 for LTVs. In spite of these differences, the authors find that, overall, the groups have similar results and the differences are explained by covariates such as comorbidities.

Uses: The authors think this information will assist families and doctors in making end-of-life care decisions for ICU patients. They think the emphasis on STV versus LTV is misdirected. They recommend ventilation for fairly healthy patients with few comorbidities who need assistance while anesthetized or sedated and who have experienced a discrete acute event amenable to treatment.

Limitations: Definitions are inconsistent in the literature on LTV. In the current study, the refusal rate was high. The 538 patients tracked were selected from more than 3,000. Most failed some criteria the authors established but almost 30% of the families refused.

Esserman, L., H. Cowley, et al. (2002). "Improving the Accuracy of Mammography: Volume and Outcome Relationships." Journal of the National Cancer Institute **94**: 369-375.

Keywords: Mammography, quality, volume

Purpose: The ability of British screeners is compared with US screeners.

Data: A standardized screener test, PERFORMS 2, was administered to 194 UK

radiologists and 60 US radiologists. The US group was subdivided into high, medium and low volume screeners.

Methods: ROC curves were fitted for each group and the sensitivity (percent correct positives) of the screeners was determined at 90% specificity (percent correct negatives).

Results: The US has a widely distributed system while the UK (and many other socialized medicine countries) has a centralized system. UK radiologists who screen for breast cancer read a minimum of 5,000 mammograms per year while US radiologists must read 480 per year. The US radiologists were considered high volume at 300 or more mammograms/month, medium at 100-299 mammograms/month and low volume (but in the study) at less than 100 mammograms/month. When achieving 90% specificity, the UK screeners would get 78.5% sensitivity. The high volume US radiologists would get 75.6% sensitivity, the medium volume screeners 70.2% sensitivity and the low volume screeners 64.8% sensitivity. Looked at another way, in order to limit false positives to 10%, the UK screeners missed 21.5% of cancer cases while US screeners missed between 24.4% and 35.2% of actual cancer cases. (The difference between UK screeners and the medium or low volume US screeners is statistically significant as is the difference between high volume and low volume US screeners.) The authors say the actual US approach sacrifices specificity for sensitivity (or says that eliminating false negatives is more important than eliminating false positives). We would rather do too much care than not enough. The UK thus has a higher cancer to biopsy yield (88%) than in the US (25%). The better mammography in the UK results in a considerable savings in biopsies while detecting cancer as effectively as the US approach.

Uses: This demonstrates the possibility of improving patient care and lowering costs by centralizing radiologists doing mammograms into larger centers where each does much higher volume.

Limitations: The selection process for the US radiologists attempted to get the most active screener in each clinic. Thus, these results probably overstate US average quality. The US portion of the study was done in California and may not represent other states.

Ferris, T., D. Dougherty, et al. (2001). "A Report Card on Quality Improvement for Children's Health Care." *Pediatrics* **107**: 143-155.

Keywords: Guidelines, incentives, pediatrics, quality improvement

Purpose: The authors advocate the study of children's healthcare quality improvement as a separate research field. They assess the status of this research.

Data: From more than 2000 articles mentioning quality and children's healthcare published between 1985 and 1997, they select 68 research reports. These met the criteria of studies involving a quality improvement intervention that would reduce overuse, underuse, or misuse of a clinical procedure.

Methods: This is an extensive literature review. They also interviewed a number of individuals involved in quality improvement.

Results: There was a dramatic increase in these studies. Half the studies published between 1985 and 1997 were published in the last three years. Better studies (randomized control trials) were more likely to be externally funded and reported

in one of three journals: *Pediatrics*, *Archives of Pediatrics*, or *Adolescent Care*. About one-fourth concerned guidelines or methods to increase compliance. Some involved reminder systems for vaccinations or followup. There were a number of disease management studies. The authors discuss asthma and cystic fibrosis. Contrary to the popularity of opinion leader studies in general medicine, only one study of pediatric opinion leaders was found. No study of financial incentives on pediatric practice was found. Interviews revealed that children's health care quality improvement efforts, at general hospitals, was usually limited to immunization compliance. Children's hospitals were as likely as general hospitals to have quality initiatives. These are mostly focused on cost reduction. A major lack was quality initiatives for adolescent mental health programs. Barriers to quality included lack of resources, conflict between clinical and administrative staff, organizational instability, and lack of cooperative skills among physicians. One distinct problem with children's quality efforts is the rare occurrence of negative outcomes. This hides poor quality even from the provider. Another is that it is a low cost area. So long as quality improvement is justified by cost containment, other areas than pediatrics will be targeted.

Uses: The authors state their goal of encouraging and challenging those involved in children's health care quality improvement.

Limitations: Their classifications of literature are unique.

Glance, L., T. Osler, et al. (2002). "Rating the quality of intensive care units: Is it a function of the intensive care unit scoring system?" Critical Care Medicine **30**(9): 1976-1982.

Keywords: ICU, quality measurement, risk adjustment

Purpose: Glance tests the construct validity of Intensive Care Unit (ICU) quality measurement based on risk-adjusted mortality.

Data: Sixteen-thousand patients in 32 ICUs are studied.

Methods: Several tests are done but the key was to calculate actual-to-expected mortality ratios for each ICU using three different risk-adjustment programs. The rankings of the ICU mortality results, especially identification of outliers, is compared for the three programs. Other sophisticated statistical techniques are used.

Results: The three programs' results are reasonably consistent indicating construct validity. The C (area under the ROC curve) exceeds eighty percent for each model indicating good discrimination for each. All three programs indicate that these hospitals' ICUs generally have high quality. Glance thus finds concern in the calibration of all three models. He thinks changes in quality during the period used for calibration may have had considerable impact. He notes other studies that also find mortality prediction is poor on external data sets. He notes the lack of a gold standard for risk-adjustment. On the other hand, he notes that quality improvement has been shown to affect model calibration.

Uses: This gives more comfort than prior studies that risk-adjusted mortality can indicate quality of ICUs.

Limitations: Only a third of the data could be used since it was adequate for all three scoring models.

Goldstein, S. and S. Schweikhart (2002). "Empirical Support for the Baldrige Award

Framework in U.S. Hospitals." Health Care Management Review **27**(1): 62-75.

Keywords: Baldrige, hospitals, quality improvement, TQM

Purpose: The authors test whether the health care Baldrige Criteria self-assessment of leadership, processes, systems, etc. actually leads to improved quality results for patient satisfaction, outcomes, efficiency, etc.

Data: Two hundred twenty hospitals' quality officers are surveyed concerning quality programs and results.

Methods: Outcomes, patient satisfaction, efficiency, profits, etc. are dependent variables. Independent variables are Baldrige-style indicators of quality leadership, processes, measurement, etc.

Results: A history of the Baldrige Criteria and attempts to apply it to health care and modify it for health care is provided. Other articles studying Baldrige case studies are reviewed. Several articles reviewing Total Quality Management (TQM) applications in health care providers are also reviewed. The idea has been popular in health organizations as indicated by the tens of thousands of self-assessment kits distributed. There have been no healthcare Baldrige Award winners (as of this article's publication.) The Baldrige Criteria are strong predictors of patient satisfaction. They are weak predictors of clinical outcomes and functional performance. This disappointing result was attributed to physician autonomy, risk-adjustment problems, and possible failure of the Baldrige concept. The latter could be due to its industrial origin. Are clinical outcomes really equivalent to any industrial measures and are they aggregated appropriately? Financial and market results are also weakly predicted by Baldrige Criteria. It is acknowledged that TQM applications in industry are not necessarily correlated with financial performance. The Baldrige Criteria are predictors of quality staff and internal service.

Uses: This study gives weak support for using the Baldrige criteria as a quality management tool.

Limitations: This is a cross-sectional study so causation is weak. A longitudinal study could better demonstrate that Baldrige-style quality improvement programs achieve these results. Only a fourth of those mailed a survey responded. Bias could be a result. This style of self-reported survey has been tested and found valid by others. The study is specifically relevant to hospitals. Other providers should use the results with greater caution.

Hargraves, J. L. and S. Trude (2002). "From The Field; Obstacles To Employers' Pursuit Of Health Care Quality; Inadequate information has left quality out of many employers' health care purchasing equation." Health Affairs **21**(5).

Keywords: Employer, information, purchaser, quality

Purpose: The authors study the effort and impact of employers on health quality in their communities.

Data: The Community Tracking System data is used. This is a biennial site visit/interview study of the health care situation in twelve communities.

Methods: This is descriptive analysis of the 65 interviews with some of the largest employers in each community.

Results: The employers say they are frustrated by lack of quality information on

providers. They want provider rather than plan data due to plan overlap. They would also like to see physicians' ratings of health plans. Cost remains the decisive factor in purchasing decisions. Employers were aware of safety problems. Many were aware of the Leapfrog Group effort but none of those interviewed were participating. Employer coalitions for purchasing or quality improvement had been formed in each of the twelve markets but had disbanded in six. The decline in influence was due to organized opposition of the providers, lack of perceived benefit to the largest employers, and inconsistent goals. In spite of the above, value-based purchasing ideas have had an impact. Almost all of the employers interviewed used an RFP to solicit health plan proposals and quality standards were often included. Consultants usually handled the bid request/selection process. The high cost of changing plans was a concern. Some employers were confident of their local providers quality and saw little benefit from quality efforts. The largest employers were national firms with employees spread around the US so they thought their quality efforts should be at a national level. There was an increase in the attempt to make employees better purchasers of healthcare. Many used personal spending accounts. They provided employees consumer information such as access to websites with provider quality information.

Uses: This analyzes the status and trends of local quality efforts in an effort to inform policy.

Limitations: The authors note piecemeal progress.

Jencks, S. F., T. Cuerdon, et al. (2000). "Quality of Medical Care Delivered to Medicare Beneficiaries: A Profile at State and National Levels." Journal of the American Medical Association **284**(13): 1670-1676.

Keywords: Compliance, fee for service, measurement, Medicare, quality

Purpose: The authors set out to determine compliance with established guidelines for care in Medicare fee-for-service (FFS) settings. They claim to fill a void since managed care quality has been extensively studied. Compliance within each state was another study subject. This is a Health Care Financing Administration (HCFA, now called Center for Medicare/Medicaid Services, CMS) funded project partly fulfilling its congressional responsibility to monitor and promote quality.

Data: For each of 24 established, accepted guidelines, over 700 cases were drawn in each state. Compliance, or whether some factor rationalized noncompliance, was determined for each. The authors list five reasons why process is a better measure than outcomes for this type of project.

Methods: The data analysis was extensive. The end result was to calculate the conformance in each state on each guideline. Overall results for each guideline and state were determined.

Results: The median state's compliance on each guideline ranged from 11% to 95%. Each states' ranks on the 24 items were averaged and the results ranged from 10 to 48. The Northern and less populous states scored better than the Southern and more populous states. However, the overall level of compliance, 69%, is the real result. The authors expect that many lives could be saved by full compliance but find the numerical estimate beyond their scope.

Uses: The state-by-state information is interesting but poor overall compliance is the

real message. Medicare intends to be a purchaser of quality health care. The authors also conclude that better quality health care requires a government-provider-plan-beneficiary partnership.

Limitations: The authors give two pages of Qualifications but they do note some consistency with prior studies of managed care. Their choice of the 24 (out of thousands) guidelines is discussed. They discuss the possibility that there were more cases in which the guidelines should have been skipped.

Kiefe, C. I., J. J. Allison, et al. (2001). "Improving Quality Improvement Using Achievable Benchmarks For Physician Feedback: A Randomized Controlled Trial." The Journal of the American Medical Association **285**(22): 2871-2879.

Keywords: Benchmarks, Medicare, quality improvement, quality measures

Purpose: The authors express two purposes. First they demonstrate how a randomized control trial (RCT) may be used to test the efficacy of a quality improvement methodology. Second, they test a particular quality improvement process called Achievable Benchmarks of Care (ABC).

Data: Seventy Alabama physicians treating 2978 fee-for-service Medicare patients with diabetes are studied. This is a study within a study. Under the Health Care Financing Administration (HCFA-now called Center for Medicare/Medicaid Services or CMS) many quality improvement projects are being funded including the Ambulatory Care Quality Improvement Project (ACQIP). ACQIP is studying different approaches to improving the quality of physicians' treatment of diabetes. ABC is one of several methods studied in Alabama.

Methods: ACQIP generally tried to test methods of improving guideline compliance by using feedback and education, so ABC is a natural fit. CMS defined the quality process measures. They included testing for glucose control; screening for cholesterol, triglycerides, and creatinine; and doing foot examinations and influenza vaccinations. ACQIP regularly informed all physicians of their performance and the mean of their peers. ABC only added an indicator for best of peers. Other quality improvement techniques tested in ACQIP included group meetings, chart interventions, patient educational material, reminders, standing orders, root cause analysis, etc. Only ABC was subject to this RCT. The authors give references to other articles on ABC and do not give much description here. Both the control and the experimental group improved in each test. (Remember the control group was given feedback against average performance of peers.) In nearly all tests the ABC intervention showed a clear advantage over the control.

Uses: First, a method for using RCT in a quality improvement project is demonstrated. Secondly, ABC is shown to have a positive impact on compliance with guidelines.

Limitations: The authors are candid in listing a number of qualifiers to their study. One concern, which they say is the subject of much current research, is whether quality achieved through feedback will persist over time. The participating physicians were volunteers. There are some data requirements that may not always be available. The ABC was added to the fairly intense ACQIP. Would it work as well in a less intense quality environment?

Kohn, L. T., J. M. Corrigan, et al. (2000). "To Err is Human: Building a Safer Health System." 287.

Keywords: Medical errors, patient safety, quality

Purpose: We review the Executive summary of this 287 page report. The editors' stated goal is to shake things up. Patient safety is a serious problem that is not getting the attention it deserves.

Data/Methods: The book builds on extensive earlier research and academic reporting. It is a synthesis of the state of health care quality research in 2000.

Results: The authors quote estimates that between 44,000 and 98,000 die in hospital accidents each year. The financial and economic consequences are described.

These estimates are likely understated. Much less serious problems get public and congressional attention but, generally there is silence about medical errors.

Recommendations: 1. Aviation, the workplace, and the environment are much safer today than 30-50 years ago. Some of the credit must go to the FAA, OSHA, and EPA. A medical care equivalent is needed to provide leadership and focus.

2. Quality improvement depends heavily on learning from errors. A new norm of error reporting, analysis, and future prevention through improved systems must be established. The current system of finding someone to blame encourages hiding errors. 3. Improved standards and guidelines must continue to be developed. They must be disseminated much more effectively. They must place more focus on patient safety. 4. Quality improvement must be emphasized in every provider.

Uses: This was an effective call-to-arms for quality improvement.

Limitations: This is not a how-to book. It states principles which must be applied in many areas with much more work.

Krauss, J. and R. Maclean (2002). "Inpatient Mortality: A Reflection of Quality Care?" Outcomes Management **6**(4).

Keywords: Mortality, quality measures, risk adjustment

Purpose: The authors challenge the common practice of using mortality results as an indicator of health care quality.

Data/Methods: Not applicable.

Results: The authors assert the need for risk adjustment but then question whether it can be adequately performed. These are not new ideas. They attribute some of their criticism of mortality as a quality indicator to Florence Nightingale. She also recognized the need for risk adjustment. They note the valid concern that risk adjustment works well for large groups but does not discriminate between small groups or individuals. They also note that measures of process may be more important than mortality results especially for the many procedures with low mortality expectation.

Uses: This is passionate ammunition for critics of mortality-based quality measures and a wake-up call and warning to those who provide or use such measures.

Limitations: No data is used. A clear understanding of risk adjustment is not evident.

Krumholz, H., S. Rathore, et al. (2002). "Evaluation of a Consumer-Oriented Internet Health Care Report Card: The Risk of Quality Ratings Based on Mortality Data." The Journal of the American Medical Association **287**(10): 1277-1287.

Keywords: Information, Internet, quality ratings

Purpose: The authors review Healthgrades.com, the largest Internet source of provider quality ratings.

Data: Acute myocardial infarction (AMI) ratings of more than 200,000 hospitals are analyzed. Healthgrades' ratings are taken from their website and compared with data from the Cooperative Cardiovascular Project. The latter is from the Centers for Medicare and Medicaid Services.

Methods: Healthgrades rates hospitals (and other providers) one-star through five-star (best). The authors compared the overall mortality and quality process results of each ratings group of hospitals. Further, they compared the ability to discriminate between individual hospitals by comparing every pair of hospitals from different ratings groups to see if they were significantly different.

Results: The results sound like the usual risk-adjustment research results. The overall mortality and quality results of each class is significantly different but the ability to discriminate between individual hospitals is very poor. The one-star hospital group had a 22% thirty-day mortality rate while five star hospital group had a 16% thirty-day mortality rate. Other classes were in between and in the expected order. However, in random pairs of hospitals, one chosen from the one-star group and the other chosen from the five-star group, the difference between the two hospitals was statistically significant in only 3% of the pairs. The authors think this kind of data is useful for providing direction in quality improvement but not for its intended purpose of helping consumers make choices about where to receive care. The website's description of the meaning of the ratings leads to further confusion.

Uses: The authors begin the job of quality rating of the health care quality raters. Many such services exist and their quality is unknown. Healthgrades, like most other services, says its methodology is proprietary. Thus this indirect analysis of their results is necessary.

Limitations: The authors had to make many decisions about the data they would include or discard and about their methods. These are subject to criticism which is a key point. Healthgrades' methods cannot be criticized since they are a black box. Actually, Healthgrades does reveal that it uses administrative data. The authors have used better clinical data.

Landon, B., A. Zaslavsky, et al. (2001). "Health Plan Characteristics And Consumers' Assessments Of Quality; For the first time, the characteristics of health plans are linked with consumer feedback in a nationwide survey." Health Affairs **20**(2).

Keywords: Consumer satisfaction, HMO, Medicare

Purpose: The authors seek to determine if consumer satisfaction with their HMO is a function of HMO plan characteristics.

Data: Survey responses from 82,000 Medicare beneficiaries from 182 health plans are reviewed. (HCFA's CAHPS study.)

Methods: Consumers' assessments of their care is associated with various plan variables.

Results: Plan variables are its age, type, profitability, geographic scale, location (region), size, federal qualification, Medicaid dummy, POS dummy, PPO dummy, and NCQA status. The dependent variables that got the most attention

in this study were plan rating, access, and customer service. Plan size, model type, and accreditation status were not good predictors. Region was a strong predictor with lowest scores in the Pacific region and highest in the Northeast and North MidAtlantic. Tax status and national scale were strong predictors of performance. For-profit plans had worse results on access and customer service. National plans did not do as well as local plans. The for-profit plans are mostly in the Pacific region so confounding was studied and for-profit plans were still determined to underperform. Actual higher disenrollment rates were consistent with poorer ratings on plan satisfaction. The authors say their results are consistent with others on for-profit performance.

Uses: The authors feel this information can help set guidelines for plan accreditation.

Limitations: The authors acknowledge looking at only a few plan characteristics on which data was easily accessible. Other characteristics could be pertinent and confound these results. The data is cross-sectional. Causation is harder to accept. The study population is Medicare. It may not pertain to employee groups.

Langemo, D., J. Anderson, et al. (2002). "Nursing Quality Outcome Indicators: The North Dakota Study." The Journal of Nursing Administration **32**(2): 98-105.

Keywords: Nursing, outcomes, quality

Purpose: The authors state their goals as wanting to prove the need for more and better-trained nurses.

Data: This is from a pilot study done in North Dakota in preparation for a national study. Two-hundred nurses and 900 patients are surveyed. Other data on certain patient care outcomes such as pressure ulcers and falls is also obtained.

Methods: Outcomes were related to the number of nurses, experience, and licensing levels of nurses.

Results: Other studies are quoted as generally showing better quality outcomes with more and better trained nurses. These range from patient satisfaction studies to risk-adjusted mortality results. The authors assert that hospitals have greatly increased the number of patients per nurse and have shifted duties to lower licensed nurses. The results of this study were mixed as regards their stated objective. Falls seemed to respond to the number rather than the quality of nurses on duty. Pressure ulcers were reduced by higher quality nurses. Medication errors were not studied but would seem the real benefit of better training. Patient satisfaction with nursing and the hospitals in general was high. Nurse job satisfaction was low. It depended highly on workload and the educational level of coworkers and supervisors.

Uses: The authors state that they are providing useful management information.

Limitations: Although data analysis was done by a research firm, the rest of the study seems weak, perhaps based on its admittedly biased purpose.

Lansky, D. (2002). "Improving Quality Through Public Disclosure Of Performance Information; A refreshingly simple idea that uses government authority to force an issue that has not made much progress with market power alone." Health Affairs **21**(4).

Keywords: Evaluation, information, outcomes, quality

Purpose: This provides theory, history, and advocacy of public reporting of quality

results especially from government-funded programs.

Data/Methods: Not applicable.

Results: Justice Brandeis wrote that "electric light is the most efficient policeman" and Justice Breyer said "disclosure provides an intermediate means of addressing public risks without creating impossibly complex regulatory systems or reducing the beneficial affects of choice in the marketplace." Further, it is good public policy that any government expense require demonstration that the expense is achieving its goal. He mentions the New York cardiology program in which government requirements for reporting led to great improvement. The record of government expense accountability in healthcare is spotty but improving. Only 0.1% of the Medicare budget (for peer review organizations) is directed towards assurance that the expense meets its purpose. Only 1% of the National Institute of Health (NIH) budget (for the Agency for Healthcare Research and Quality) is directed at seeing that medical expense actually achieves the desired outcome. TANF and HIPAA increased federal spending on healthcare without increasing reporting. Lansky says the State Children's Health Insurance Program (SCHIP) is a model for future expense. It required the states to report on the effectiveness and quality of the increased medical care provided. This led to the Child and Adolescent Health Measurement Initiative (CAHMI), a program initially designed to help states meet the administrative burden. CAHMI has blossomed into a model for development of quality and organizational effectiveness measurement tools.

Uses: Lansky recommends a similar approach for any further government investment in healthcare such as Medicare drug benefits or tax credits. He says studies indicate that 25% of prescriptions now involve errors by the physician or pharmacist. Further, the patients don't follow directions. Congress should not invest billions knowing that more than 25% will be wasted or cause great harm. Accountability for the expense should be required.

Limitations: He acknowledges the need for cost-benefit analysis of regulations but doesn't suggest an approach.

McGlynn, E. and R. Brook (2001). "Keeping Quality On The Policy Agenda; How many more people have to die before we accept that quality is everyone's problem?" Health Affairs **20**(3).

Keywords: Information, public awareness, quality improvement

Purpose: Frustrated by the lack of Congressional and public concern, the authors call for a consistent quality improvement program for US health care.

Data/Methods: Not applicable.

Results: The authors make much of the Firestone tire situation. The Congress and public demanded that this life-threatening quality assurance problem be solved. During the period of the tire failures, "a few" (20?) were killed annually. The Institute of Medicine estimates that 98,000 are killed annually by medical errors. The public isn't concerned and Congress won't deal with it. The authors blame diffuse responsibility, cognitive dissonance, nineteenth-century systems, lack of information, and a tendency to shoot the messenger. They call for leaders or champions of quality. They think the various private organizations that help fund health care research should be more aware of quality. For instance, the National

Cancer Society (NCS) helps researchers develop a new drug. If the doctors and pharmacists proceed to misuse it in 25% of cases causing its ineffectiveness or deaths, shouldn't the NCS be concerned? The authors call for government funding of systems, reporting, and research for quality.

Uses: Advocacy.

Limitations: The authors veered from the intended purpose of discussing how to keep quality on the public and Congressional agenda.

Miller, R. and H. Luft (2002). "HMO Plan Performance Update: An Analysis Of The Literature, 1997-2001; By and large, HMOs have not accomplished what their proponents promised: changing clinical practice processes and improving quality, while lowering costs for both purchasers and consumers." Health Affairs **21**(4).

Keywords: HMO, performance measures, quality

Purpose: This reviews literature on HMO performance and synthesizes their results, especially regarding quality.

Data: Out of hundreds of articles initially screened, the authors chose 79 as adequately testing a result about the relative performance of HMO and other health plans.

Methods: They carefully categorize and summarize the articles' results. For each category they count the articles as favoring HMOs, nonHMOs or neutral.

Results: This is an update to two prior similar studies the authors did. This time they increased the focus on quality. They found more articles than previously indicating more researchers' interest, which is of course driven by more funding. People want to know the impact of HMOs. HMOs and nonHMOs have mixed results as regards quality. HMOs use fewer resources. HMOs cause access problems and achieve poorer satisfaction scores. HMOs do better on prevention efforts. HMOs have a similar number of hospital admissions but the stays are shorter. Some of these effects pass through to their communities. (HMO market penetration is associated with poorer access, lower costs, and better prevention.)

Uses: This is good quality information for informing public policy and HMO management. The authors say HMOs need better systems and incentives to fulfill their promise.

Limitations: The authors discuss research selection bias and publication bias. There were tradeoffs in article selection and scoring but the authors tried to be very candid about their approach. They question if risk adjustment still favors HMOs by paying them too much for their insureds thus improving their quality.

Newhouse, J. (2002). "Why Is There A Quality Chasm? ; The barriers to good performance are more fundamental than simply the lack of organized systems." Health Affairs **21**(4).

Keywords: Economics, fundamentals, incentives, quality

Purpose: This gives a rather pessimistic view that errors and poor quality will remain a part of the health care delivery system. The systems approach advocated in the IOM Quality Chasm report cannot achieve industrial level quality.

Data/Methods: Not applicable.

Results: Newhouse gives the economic view generally along the lines Arrow laid out 40

years ago. Medical provision will always be filled with uncertainty. Risk-averse people will therefore buy insurance. There is asymmetric information between doctor and patient. Moral hazard will cause patients to overconsume health care when insured. Doctors and hospitals will provide too much care when the patient desires it. The large public sector involvement provides incentives other than efficiency. Rapid technological change adds to the uncertainty. HMOs will have a difficult time intervening as the patient trusts the doctor as his agent rather than the HMO. Generally, the lack of a price mechanism will not encourage efficiency. Efficiency usually means production of the required product (including quality) at the lowest cost but it can mean production of the highest quality at a fixed cost. The providers do not have incentive to be efficient. Most savings from investment in efficiency would accrue to the third party payor and actually reduce provider income.

Uses: This discourages those who think medical care could achieve industrial level quality by emulating industrial programs.

Limitations: There is little data. This is out of the mainstream of quality articles in medicine.

Rosner, F., J. Berger, et al. (2000). "Disclosure and Prevention of Medical Errors." Archives of Internal Medicine **160**(14): 2089-2092.

Keywords: Disclosure, errors, quality improvement

Purpose: The authors discuss the professional and moral obligations of physicians to disclose their own and others' errors to patients.

Data/Methods: Essay.

Results: The physician is generally, but not always, obligated to inform the patient of errors. Errors are common. Consequences range from serious to inconsequential. They can be caused by limited knowledge or experience, poor judgment, or carelessness. The latter is most common. The AMA Code 9 requires the physician to disclose errors that cause significant medical complications. The physician must balance relations with patients and peers, the patient's right to know, the possibility that awareness of the error may harm the patient, and the possibility that revealing the error is therapeutic. Discussion of bad outcomes (without error) is also important. Physicians are not trained to deal with their inevitable errors. Failure to disclose errors to peers makes the errors more likely to be repeated. The author presents a situational ethics discussion of times when nondisclosure is in the patient's best interest. Other physicians and other hospital personnel should report physicians' errors and should not be punished for doing so. Hospitals should accept any liability created by disclosure as a price paid for improvement. The author discusses the possibility that the system rather than an individual may be at fault. Revelation of errors is necessary for quality improvement.

Uses: This demonstrates how deeply the physicians feel personally accountable for medical success and failure. They will not easily give up their total responsibility in exchange for being part of a team or system that provides care.

Limitations: This is an opinion piece with a very narrow point of view.

Scalise, D. (2001). "Who's Rating You?" Hospital Health Network **75**(12): 36-40.

Keywords: Consumer information, Internet, ratings, quality

Purpose: Like other provider-oriented articles on consumer information, this is highly critical of what is available to the public. It is also an advocacy piece, saying that providers should get involved in disseminating quality information before those "not on the frontlines of care" have defined quality for the consumer.

Data/Methods: This is an opinion piece.

Results: A number of consumer health information websites are listed and briefly described. This is one of the most useful aspects of the article. Otherwise, it says the data is generally inconsistent, the methods are secret or poor, and the assumptions are misplaced. Scalise acknowledges that consumers desire quality information and that the amount available on the web will only grow. The complexity of the subject is also lost on the average consumer who wants simple ratings. Statistical data is desired by regulators, employers, and plans but not by individuals.

Uses: This helps us understand the concern of providers about published quality information. The website list may be useful.

Limitations: This is not from academic or reviewed literature.

Shannon, S., P. Mitchell, et al. (2002). "Patients, Nurses, and Physicians Have Differing Views of Quality of Critical Care." Journal of Nursing Scholarship **34**(2): 173-179.

Keywords: ICU, patient satisfaction, quality

Purpose: The authors compare the views of patients, physicians, and nurses regarding quality care in the critical care unit (ICU).

Data: 489 patients, 518 nurses, and 515 physicians from 25 ICUs in 14 Pacific Northwest hospitals were interviewed.

Methods: Data were aggregated and analyzed within each ICU to compare quality opinions. Instruments were common research tools validated in earlier studies.

Results: Physicians had the most favorable opinions of quality care and nurses the least. Physicians overestimated patient satisfaction but nurses correctly matched patient satisfaction. Some prior studies had assumed nurses or physicians could speak for the patients (who are often unable to express opinions on satisfaction) in the ICU. This confirms that nurses might do so. Factors that seemed to drive the physician and nurse opinion of quality of care were items related to their work environment and professional standards.

Uses: Encourages research on quality to seek multiple viewpoints. The patient's opinion is key and may not be the same as professionals.

Limitations: The researchers just looked at correlations between many different variables. Further statistical analysis could have contributed much.

Shi, L. (2000). "Type of Health Insurance and the Quality of Primary Care Experience." American Journal of Public Health **90**(12): 1848, 8p.

Keywords: HMO, Insurance, primary care, quality

Purpose: The authors study how the type of insurance affects primary care as viewed by the patient.

Data: Data was from the 1996 Medical Expenditure Panel Survey. Patients were under

age 65.

Methods: Logistic regression was used to see if insurance status predicted several indicators of primary care quality. Two rounds of interviews with each patient were completed. More than 10,000 patients were included. Statistical analysis was sophisticated.

Results: Other research has established the importance of primary care to health. It has also established the importance of insurance to getting access to primary care. This adds to the research by studying the influence of type of insurance. Insurance status included uninsured, fee-for-service (FFS) privately insured, managed care (MC) privately insured, and publicly insured. Indicators of primary care are first contact, longitudinality, comprehensiveness, and coordination. These were used in earlier research. Ten questions related to the first contact and generally determined how easy it was to get the initial appointment for a new illness. Six questions tested the continuity of care and patient-physician-relationship. Only three questions dealt with comprehensiveness and coordination. Many other demographic covariates were included to see how they might affect the insurance-primary care relationship. Covariates reflecting who might have the greatest need were also tested. Overall, the uninsured fared the worst. Those privately insured did better than those publicly insured. FFS did better than MC at least as regards access and longitudinal care. Those with greatest need were most likely to be publicly insured. Insurance status has more impact on the first contact than on continuing care. Those in MC did better on preventative care but worse on all measures of continuing care than those in FFS. Those in MC were more likely to name a facility as their primary care physician while those in FFS named an individual. Those in FFS had fewer changes of physician, more communication with, trust in, and satisfaction with their physician, and better coordination of specialists.

Uses: This should inform public policy on the expansion of insurance to the uninsured. Public programs are not achieving the care of private insurance. The expansion of MC should be monitored carefully.

Limitations: This is comprehensive. It is also likely to change with time and needs updating. Comprehensiveness and coordination measures were more limited. The varieties of MC could be addressed in future research. This tested insurance status against process of care. Insurance status and outcomes could be studied. The follow-up interviews were not used in a longitudinal way. This is cross-sectional analysis so causal inference is weaker.

Sprague, L. (2001). "Quality in the Making." The American Journal of Medicine **111**(5): 422-431.

Keywords: Quality improvement, quality programs

Purpose: Sprague provides an overview of quality assurance and improvement in US health care.

Data/Methods: This is an opinion piece.

Results: She says there is no agreed-upon frame work for understanding or improving health care. One widely used approach is that of Donabedian and she follows it in reviewing structure, process, and outcomes. She looks at the roles for patients, providers, and regulators in requiring quality. She provides one of the

simpler guides to most of the alphabet soup of players in health care quality. The Agency for Healthcare Research and Quality (AHRQ) is a federal entity which encourages (funds) research and private initiatives and also maintains several important health care data sets. It is especially interested in publishing guidelines for best practices. Three organizations accredit health (managed care) plans. One of them, the National Committee for Quality Assurance, maintains process measures in its Health Plan Employer Data and Information Set (HEDIS). It is the source of data for much research. A major outcome measurement is patient satisfaction. The Center for Medicare and Medicaid Services (CMS), AHRQ, and others sponsor the Consumer Assessment of Health Plans Survey (CAHPS) to track patient satisfaction. Leapfrog is a business-oriented group that wants to bring market (employer) pressure to bear on provider quality improvement. It does have government affiliation with CMS and the Office of Personnel Management (OPM). Some state governments provide health consumer information, mostly on hospitals. Besides its role as regulator, the government is the largest purchaser of health care. Besides CMS and OPM, the Departments of Defense and Veterans Affairs provide healthcare. Together, they have formed the Quality Interagency Coordination Task Force (QuIC). CMS has always played a role in quality through Peer Review Organizations (PROs) but the quality role may be compromised by the cost cutting role. In 1992, CMS and the PROs started the Health Care Quality Improvement Program (HCQIP) to encourage best practices. CMS and PROs are also working to eliminate racial disparities in care. CMS specifically studies managed care quality in its Quality Improvement System for Managed Care (QISMC). Sprague also discusses some public-private efforts such as the National Forum for Health Care Quality Measurement (NQF). Its goal is to measure and report quality. Several private groups report quality information on the Internet, assuming that an informed consumer market will drive quality improvement. CMS is involved here as well with its Center for Beneficiary Services, an educational effort.

Uses: This is an excellent paper with which to begin reading about US health care quality initiatives.

Limitations: As with any overview, there is minimal depth in the descriptions.

Tu, H. T. and J. D. Rechovsky (2002). "Assessments of Medical Care by Enrollees in For-Profit and Nonprofit Health Maintenance Organizations." The New England Journal of Medicine **346**(17): 1288-1293.

Keywords: HMO, nonprofit, patient satisfaction

Purpose: Patient satisfaction with care is compared and contrasted in for profit and nonprofit HMOs.

Data: Extensive data from two national databases is used. 12,000 records are matched. The sources are the Community Tracking Study: the Household Survey and the Insurance Followback Survey.

Methods: Sophisticated techniques are used to determine the effect of plan ownership status on various satisfaction questions controlling for several participant and plan parameters. Generally, multivariate logistic regression is the main tool. Plan characteristics that required control to avoid confounding included age as the nonprofits were much older and scope as the for-profits were more likely to be

national while the nonprofits were more local.

Results: On many criteria, and certainly in the overall results, there is little to choose between HMOs with ownership differences. The one significant exception was for the fewer than 1,000 patients who assessed themselves as unhealthy. Those in nonprofit plans were about ten percent more likely to rate their care excellent than those in for-profit plans in these categories: overall satisfaction, physician thoroughness, explanations, and listening.

Uses: With for-profit plans growing much faster than nonprofits some have feared that care will decline. This generally refutes the concern but shows the need to monitor the situation of the less healthy in for-profit HMOs.

Limitations: The situation evolves quickly so this 1996-1998 data will soon be outdated. In many ways, market and regulatory forces cause the difference between the HMOs to diminish.

WHI, W. s. H. I. (2002). "Risks and Benefits of Estrogen Plus Progestin in Healthy Postmenopausal Women: Principal Results From the Women's Health Initiative Randomized Control Trial." Journal of the American Medical Association **288**(3): 321-333.

Keywords: Hormone treatment, outcomes, quality

Purpose: This is a major study to determine the positive and negative impacts of "the most common hormone treatment in the US". The expectation was that a positive effect on coronary heart disease (CHD) might be offset by a negative effect on breast cancer. Other conditions were also monitored. It is part of a much larger project to improve the quality of women's health care.

Data: Over 16,000 women were recruited nationally at 40 clinics for an eight year study with randomized assignment to treatment and placebo groups.

Methods: The women were contacted each six months and given an annual physical exam. The impact of the treatment was studied using various survival model techniques: Kaplan Meier, hazard ratios, and Cox proportional hazards.

Results: The trial was stopped during the fifth year when the excess breast cancers went out of bounds and there was actually an increase in CHD. Strokes and pulmonary embolisms were also high for the treatment group while other cancers and hip fractures were low. Overall mortality was about the same in each group (but there had been few deaths from cancer to that point-just cases. A "global index" representing all the risks under study was 15% higher for the treatment group.

Uses: Hormone replacement therapy should not be recommended for CHD effects in healthy postmenopausal women.

Limitations: Impacts on menopause side effects were not studied. Only a particular dose taken orally was studied. Other treatment modalities might fair better. The early stopping of the study could have masked long term benefits that were slow to materialize but it was clear from the trends that any overall positive benefit would not show within the eight year study design. Results may be underestimated due to some participants leaving the study. The separate effects of estrogen and progestin are not studied.