Chapter 8: The Safety and Quality of Health Care

David W. Bates

INTRODUCTION

Safety and quality are two of the central dimensions of health care. In recent years, it has become easier to measure safety and quality, and it is increasingly clear that performance in both dimensions could be much better. The public is—with good justification—demanding measurement and accountability, and payment for services will increasingly be based on performance in these areas. Thus, physicians must learn about these two domains, how they can be improved, and the relative strengths and limitations of the current ability to measure them.

Safety and quality are closely related but do not completely overlap. The Institute of Medicine has suggested in a seminal series of reports that safety is the first part of quality and that the health care system must first and foremost guarantee that it will deliver safe care, although quality is also pivotal. In the end, it is likely that more net clinical benefit will be derived from improving quality than from improving safety, though both are important and safety is in many ways more tangible to the public. The first section of this chapter will address issues relating to the safety of care and the second will cover quality of care.
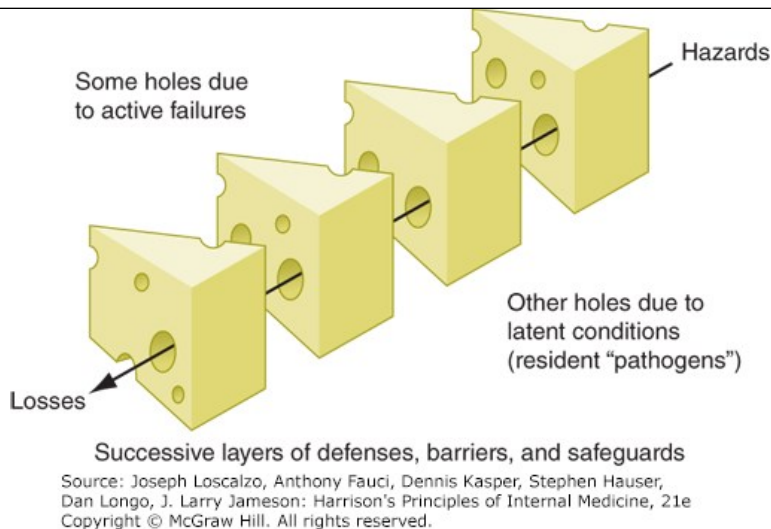
SAFETY IN HEALTH CARE

Safety Theory and Systems Theory

Safety theory clearly points out that individuals make errors all the time. Think of driving home from the hospital: you intend to stop and pick up a quart of milk on the way home but find yourself entering your driveway without realizing how you got there. Everybody uses low-level, semiautomatic behavior for many activities in daily life; this kind of error is called a *slip*. Slips occur often during care delivery—e.g., when people intend to write an order but forget because they must complete another action first. *Mistakes,* by contrast, are errors of a higher level; they occur in new or nonstereotypic situations in which conscious decisions are being made. An example would be dosing of a medication with which a physician is not familiar. The strategies used to prevent slips and mistakes are often different.

Systems theory suggests that most accidents occur as the result of a series of small failures that happen to line up in an individual instance so that an accident can occur (Fig. 8-1). It also suggests that most individuals in an industry such as health care are trying to do the right thing (e.g., deliver safe care) and that most accidents thus result from defects in systems. Systems should be designed both to make errors less likely and to identify those that do inevitably occur.

FIGURE 8-1

"Swiss cheese" diagram. Reason argues that most accidents occur when a series of "latent failures" are present in a system and happen to line up in a given instance, resulting in an accident. Examples of latent failures in the case of a fall might be that the unit is unusually busy and the floor happens to be wet. *(Adapted from J Reason: BMJ 320:768, 2000.)*

Source: Joseph Loscalzo, Anthony Fauci, Dennis Kasper, Stephen Hauser,
Dan Longo, J. Larry Jameson: Harrison's Principles of Internal Medicine, 21e
Copyright © McGraw Hill. All rights reserved.

Factors That Increase the Likelihood of Errors

Many factors ubiquitous in health care systems can increase the likelihood of errors, including fatigue, stress, interruptions, complexity, and transitions. The effects of fatigue in other industries are clear, but its effects in health care have been more controversial until recently. For example, the accident rate among truck drivers increases dramatically if they work over a certain number of hours in a week, especially with prolonged shifts. A recent study of house officers in the intensive care unit demonstrated that they were about one-third more likely to make errors when they were on a 24-h shift than when they were on a schedule that allowed them to sleep 8 h the previous night. The American College of Graduate Medical Education has moved to address this issue by putting in place the 80-h workweek. Although this stipulation is a step forward, it does not address the most important cause of fatigue-related errors: extended-duty shifts. High levels of stress and heavy workloads also can increase error rates. Thus, in extremely high-pressure situations, such as cardiac arrests, errors are more likely to occur. Strategies such as using protocols in these settings can be helpful, as can simple recognition that the situation is stressful.

Interruptions also increase the likelihood of error and occur frequently in health care delivery. It is common to forget to complete an action when one is interrupted partway through it by a page, for example. Approaches that may be helpful in this area include minimizing interruptions and setting up tools that help define the urgency of an interruption.

Complexity represents a key issue that contributes to errors. Providers are confronted by streams of data (e.g., laboratory tests and vital signs), many of which provide little useful information but some of which are important and require action or suggest a specific diagnosis. Tools that emphasize specific abnormalities or combinations of abnormalities may be helpful in this area.

Transitions between providers and settings are also common in health care, especially with the advent of the 80-h workweek, and generally represent points of vulnerability. Tools that provide structure in exchanging information—for example, when transferring care between providers—may be helpful.

The Frequency of Adverse Events in Health Care

Most large studies focusing on the frequency and consequences of adverse events have been performed in the inpatient setting; some data are available for nursing homes, but much less information is available about the outpatient setting. The Harvard Medical Practice Study, one of the largest studies to address this issue, was performed with hospitalized patients in New York. The primary outcome was the adverse event: an injury caused by medical management rather than by the patient's underlying disease. In this study, an event either resulted in death or disability at discharge or prolonged the length of hospital stay by at least 2 days. Key findings were that the adverse event rate was 3.7% and that 58% of the adverse events were considered preventable. Although New York is not representative of the United States as a whole, the study was replicated later in Colorado and Utah, where the rates were essentially similar. Since then, other studies using analogous methodologies have been performed in various developed nations, and the rates of adverse events in these countries appear to be ~10%. Rates of safety issues appear to be even higher in developing and transitional countries; thus, this is clearly an issue of global proportions.

In the Harvard Medical Practice Study, adverse drug events (ADEs) were most common, accounting for 19% of all adverse events, and were followed in frequency by wound infections (14%) and technical complications (13%). Almost half of adverse events were associated with a surgical procedure. Among nonoperative events, 37% were ADEs, 15% were diagnostic mishaps, 14% were therapeutic mishaps, 13% were procedure-related mishaps, and 5% were falls.

ADEs have been studied more than any other error category. Studies focusing specifically on ADEs have found that they appear to be much more common than was suggested by the Harvard Medical Practice Study, although most other studies use more inclusive criteria. Detection approaches in the research setting include chart review and the use of a computerized ADE monitor, a tool that explores the database and identifies signals that suggest an ADE may have occurred. Studies that use multiple approaches find more ADEs than does any individual approach, and this discrepancy suggests that the true underlying rate in the population is higher than would be identified by a single approach. About 6–10% of patients admitted to U.S. hospitals experience an ADE.

Injuries caused by drugs are also common in the outpatient setting. One study found a rate of 21 ADEs per every 100 patients per year when patients were called to assess whether they had had a problem with one of their medications. The severity level was lower than in the inpatient setting, but approximately one-third of these ADEs were preventable.

The period immediately after a patient is discharged from the hospital appears to be very risky. A recent study of patients hospitalized on a medical service found an adverse event rate of 19%; about one-third of those events were preventable, and another one-third were ameliorable (i.e., they could have been made less severe). ADEs were the single leading error category.

Prevention Strategies

Most work on strategies to prevent adverse events has targeted specific types of events in the inpatient setting, with nosocomial infections and ADEs having received the most attention. Nosocomial infection rates have been reduced greatly in intensive care settings, especially by using checklists. For ADEs, several strategies have been found to reduce the medication error rate, although it has been harder to demonstrate that they reduce the ADE rate overall, and no studies with adequate power to show a clinically meaningful reduction have been published.

Implementation of checklists to ensure that specific actions are carried out has had a major impact on rates of catheter-associated bloodstream infection and ventilator-associated pneumonia, two of the most serious complications occurring in intensive care units. The checklist concept is based on the premise that several specific actions can reduce the frequency of these issues; when these actions are all taken for every patient, the result has been an extreme reduction in the frequency of the associated complication. These practices have been disseminated across wide areas in the state of Michigan.

Computerized physician order entry (CPOE) linked with clinical decision support reduces the rate of serious medication errors, defined as those that harm someone or have the potential to do so. In one study, CPOE, even with limited decision support, decreased the serious medication error rate by 55%. CPOE can prevent medication errors by suggesting a default dose, ensuring that all orders are complete (e.g., that they include dose, route, and frequency), and checking orders for allergies, drug–drug interactions, and drug–laboratory issues. In addition, clinical decision support can suggest the right dose for a patient, tailoring it to the level of renal function and age. In one study, patients with renal insufficiency received the appropriate dose only one-third of the time without decision support, whereas that fraction increased to approximately two-thirds with decision support; moreover, with such support, patients with renal insufficiency were discharged from the hospital half a day earlier. As of 2019, over 95% of U.S. hospitals had implemented CPOE, although the decision support often is still limited.

Another technology that can improve medication safety is bar coding linked with an electronic medication administration record. Bar coding can help ensure that the right patient gets the right medication at the right time. Electronic medication administration records can make it much easier to determine what medications a patient has received. Studies to assess the impact of bar coding on medication safety are under way, and the early results are promising. Another technology to improve medication safety is "smart pumps." These pumps can be set according to which medication is being given and at what dose; the health care professional will receive a warning if too high a dose is about to be administered.

The National Safety Picture

Several organizations, including the National Quality Forum and The Joint Commission, have made recommendations for improving safety. The National Quality Forum has released recommendations to U.S. hospitals about what practices will most improve the safety of care, and all hospitals are

expected to implement these recommendations. Many of these practices arise frequently in routine care. One example is "readback," the practice of recording all verbal orders and immediately reading them back to the physician to verify the accuracy of what was heard. Another is the consistent use of standard abbreviations and dose designations; some abbreviations and dose designations are particularly prone to error (e.g., 7U may be read as 70).

Measurement of Safety

Measuring the safety of care is difficult and expensive, since adverse events are, fortunately, rare. Most hospitals rely on spontaneous reporting to identify errors and adverse events, but the sensitivity of this approach is very low, with only ~1 in 20 ADEs reported. Promising research techniques involve searching the electronic record for signals suggesting that an adverse event has occurred. These methods are not yet in wide use but will probably be used routinely in the future. Claims data have been used to identify the frequency of adverse events; this approach works much better for surgical care than for medical care and requires additional validation. The net result is that, except for a few specific types of events (e.g., falls and nosocomial infections), hospitals have little idea about the true frequency of safety issues.

Nonetheless, all providers have the responsibility to report problems with safety as they are identified. All hospitals have spontaneous reporting systems, and if providers report events as they occur, those events can serve as lessons for subsequent improvement.

Conclusions about Safety

It is abundantly clear that the safety of health care can be improved substantially. As more areas are studied closely, more problems are identified. Much more is known about the epidemiology of safety in the inpatient setting than in outpatient settings. A number of effective strategies for improving inpatient safety have been identified and are increasingly being applied. Some effective strategies are also available for the outpatient setting. Transitions appear to be especially risky. The solutions to improving care often entail the consistent use of systematic techniques such as checklists and often involve leveraging of information technology. Nevertheless, solutions will also include many other domains, such human factors techniques, team training, and a culture of safety.

QUALITY IN HEALTH CARE

Assessment of quality of care has remained somewhat elusive, although the tools for this purpose have increasingly improved. Selection of health care and measurement of its quality are components of a complex process.

Quality Theory

Donabedian has suggested that quality of care can be categorized by type of measurement into structure, process, and outcome. *Structure* refers to whether a particular characteristic is applicable in a particular setting—e.g., whether a hospital has a catheterization laboratory or whether a clinic uses an electronic health record. *Process* refers to the way care is delivered; examples of process measures are whether a Pap smear was performed at the recommended interval or whether an aspirin was given to a patient with a suspected myocardial infarction. *Outcome* refers to what happens—e.g., the mortality rate in myocardial infarction. It is important to note that good structure and process do not always result in a good outcome. For instance, a patient may present with a suspected myocardial infarction to an institution with a catheterization laboratory and receive recommended care, including aspirin, but still die because of the infarction.

Quality theory also suggests that overall quality will be improved more in the aggregate if the performance level of all providers is raised rather than if a few poor performers are identified and punished. This view suggests that systems changes are especially likely to be helpful in improving quality, since large numbers of providers may be affected simultaneously.

The theory of *continuous quality improvement* suggests that organizations should be evaluating the care they deliver on an ongoing basis and continually making small changes to improve their individual processes. This approach can be very powerful if embraced over time.

Several specific tools have been developed to help improve process performance. One of the most important is the Plan-Do-Check-Act cycle (Fig. 8-2). This approach can be used for "rapid cycle" improvement of a process—e.g., the time that elapses between a diagnosis of pneumonia and administration of antibiotics to the patient. Some statistical tools, such as control charts, are often used in conjunction to determine whether progress is being made. Because most medical care includes one or many processes, this tool is especially important for improvement.

FIGURE 8-2

Plan-Do-Check-Act cycle. This approach can be used to improve a specific process rapidly. First, planning is undertaken, and several potential improvement strategies are identified. Next, these strategies are evaluated in small "tests of change." "Checking" entails measuring whether the strategies have appeared to make a difference, and "acting" refers to acting on the results.
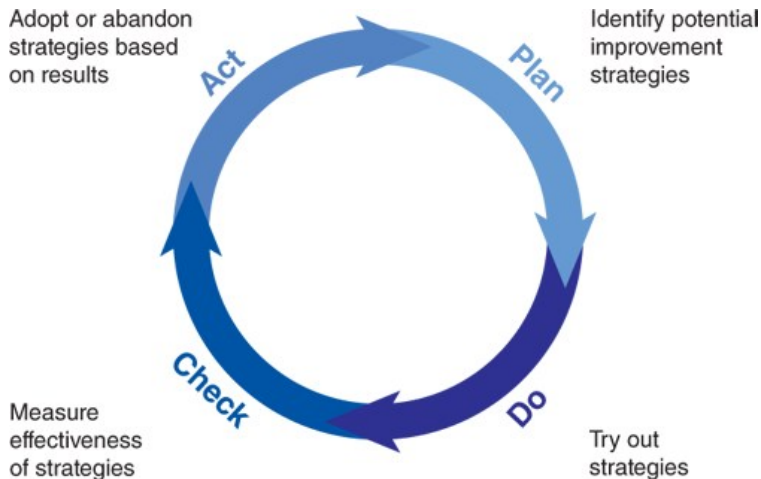
Source: Joseph Loscalzo, Anthony Fauci, Dennis Kasper, Stephen Hauser, Dan Longo, J. Larry Jameson: Harrison's Principles of Internal Medicine, 21e
Copyright © McGraw Hill. All rights reserved.

Factors Relating to Quality

Many factors can decrease the level of quality, including stress to providers, high or low levels of production pressure, and poor systems. Stress can have an adverse effect on quality because it can lead providers to omit important steps, as can a high level of production pressure. Low levels of production pressure sometimes can result in worse quality, as providers may be bored or have little experience with a specific problem. Poor systems can have a tremendous impact on quality, and even extremely dedicated providers typically cannot achieve high levels of performance if they are operating within a poor system.

Data about the Current State of Quality

A study published by the RAND Corporation in 2006 provided the most complete picture of quality of care delivered in the United States to date. The results were sobering. The authors found that, across a wide range of quality parameters, patients in the United States received only 55% of recommended care overall; there was little variation by subtype, with scores of 54% for preventive care, 54% for acute care, and 56% for care of chronic conditions. The authors concluded that, in broad terms, the chances of getting high-quality care in the United States were little better than those of winning a coin flip.

Work from the Dartmouth Atlas of Health Care evaluating geographic variation in use and quality of care demonstrates that, despite large variations in utilization, there is no positive correlation between the two variables at the regional level. An array of data demonstrate, however, that providers with larger volumes for specific conditions, especially for surgical conditions, do have better outcomes.

## Strategies for Improving Quality and Performance

Many specific strategies can be used to improve quality at the individual level, including rationing, education, feedback, incentives, and penalties. *Rationing* has been effective in some specific areas, such as persuading physicians to prescribe within a formulary, but it generally has been resisted. *Education* is effective in the short run and is necessary for changing opinions, but its effect decays fairly rapidly with time. *Feedback* on performance can be given at either the group or the individual level. Feedback is most effective if it is individualized and is given in close temporal proximity to the original events. *Incentives* can be effective, and many believe that they will prove to be a key to improving quality, especially if pay-for-performance with sufficient incentives is broadly implemented (see below). *Penalties* produce provider resentment and are rarely used in health care.

Another set of strategies for improving quality involves changing the systems of care. An example would be introducing reminders about which specific actions need to be taken at a visit for a specific patient—a strategy that has been demonstrated to improve performance in certain situations, such as the delivery of preventive services. Another approach that has been effective is the development of "bundles" or groups of quality measures that can be implemented together with a high degree of fidelity. Many hospitals have implemented a bundle for ventilator-associated pneumonia in the intensive care unit that includes five measures (e.g., ensuring that the head of the bed is elevated). These hospitals have been able to improve performance substantially. Another technique is SCAMPs, or Standardized Clinical Assessment and Management Plans. These are care guidelines developed by clinicians who identify key steps in workflow and decisions to help improve the process outcomes.

Perhaps the most pressing need is to improve the quality of care for chronic diseases. The Chronic Care Model has been developed by Wagner and colleagues (Fig. 8-3); it suggests that a combination of strategies is necessary (including self-management support, changes in delivery system design, decision support, and information systems) and that these strategies must be delivered by a practice team composed of several providers, not just a physician.

FIGURE 8-3

The Chronic Care Model, which focuses on improving care for chronic diseases, suggests that (1) delivery of high-quality care requires a range of strategies that must closely involve and engage the patient and (2) team care is essential. *(From EH Wagner et al: Eff Clin Pract 1:2, 1998.)*
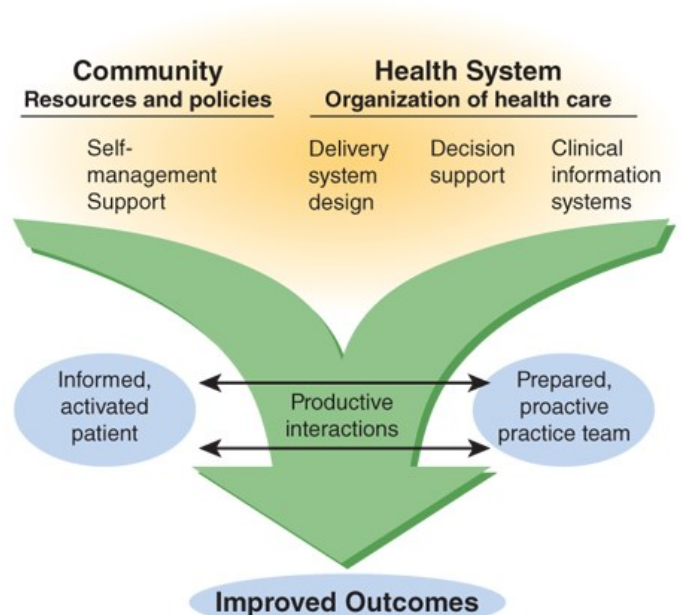
Source: Joseph Loscalzo, Anthony Fauci, Dennis Kasper, Stephen Hauser, Dan Longo, J. Larry Jameson: Harrison's Principles of Internal Medicine, 21e
Copyright © McGraw Hill. All rights reserved.

Available evidence about the relative efficacy of strategies in reducing hemoglobin A_{1c} (HbA_{1c}) in outpatient diabetes care supports this general premise. It is especially notable that the outcome was the HbA_{1c} level, as it has generally been much more difficult to improve outcome measures than process measures (such as whether HbA_{1c} was measured). In this meta-analysis, a variety of strategies were effective, but the most effective ones were the use of team changes and the use of a case manager. When cost-effectiveness is considered in addition, it appears likely that an amalgam of strategies will be needed. However, the more expensive strategies, such as the use of case managers, probably will be implemented widely only if pay-for-performance takes hold.

The evidence linking better performance on quality metrics assessing process and outcomes varies greatly by condition. For example, there is strong evidence that performing Pap smears results in better outcomes in patients who develop cervical cancer, but the evidence for many other conditions is far more tenuous.

National State of Quality Measurement

In the inpatient setting, quality measurement is now being performed by a very large proportion of hospitals for several conditions, including myocardial infarction, congestive heart failure, pneumonia, and surgical infection prevention; 20 measures are included in all. This is the result of the Hospital Quality Initiative, which represents a collaboration among many entities, including the Hospital Quality Alliance, The Joint Commission, the National Quality Forum, and the Agency for Healthcare Research and Quality. The data are housed at the Centers for Medicare and Medicaid Services, which publicly releases performance data on the measures on a website called *Hospital Compare* (*www.cms.gov/Medicare/Quality-Initiatives-Patient-Assessment-Instruments/HospitalQualityInits/HospitalCompare.html*). These data are reported voluntarily and are available for a very high proportion of the nation's hospitals. Analyses demonstrate substantial regional variation in quality and important differences among hospitals. Analyses by The Joint Commission for similar indicators reveal that performance on measures by hospitals has improved over time and that, as might be hoped, lower performers have improved more than higher performers.

Public Reporting

Overall, public reporting of quality data is becoming increasingly common. There are now commercial websites that have quality-related data for most regions of the United States, and these data can be accessed for a fee. Similarly, national data for hospitals are available. The evidence to date indicates that patients have not made much use of such data, but that the data have had an important effect on provider and organization behavior. Instead, patients have relied on provider reputation to make choices, partly because little information was available until very recently and the information that was available was not necessarily presented in ways that were easy for patients to access. Problems still exist with quality metrics; many can be "gamed," and even though providers are now nearly universally using electronic health records (EHRs), most metrics come from claims that include many inaccuracies. More metrics that leverage EHRs are sorely needed. However, many authorities think that, as more information about quality becomes available, it will become increasingly central to patients' choices about where to access care.

Pay-for-Performance

Currently, providers in the United States get paid the same amount for a specific service, regardless of the quality of care delivered. The pay-for-performance theory suggests that, if providers are paid more for higher-quality care, they will invest in strategies that enable them to deliver that care. The current key issues in the pay-for-performance debate relate to (1) how effective it is, (2) what levels of incentives are needed, and (3) what perverse consequences are produced. The evidence on effectiveness is limited, although a number of studies are ongoing. With respect to incentive levels, most quality-based performance incentives have accounted for merely 1–2% of total payment in the United States to date. In the United Kingdom, however, 40% of general practitioners' salaries have been placed at risk according to performance across a wide array of parameters; this approach has been associated with substantial improvements in reported quality performance, although it is still unclear to what extent this change represents better performance versus better reporting. The potential for perverse consequences exists with any incentive scheme. One problem is that, if incentives are tied to outcomes, there may be a tendency to transfer the sickest patients to other providers and systems. Another concern is that providers will pay too much attention to quality measures with incentives and ignore the rest of the quality picture. The validity of these concerns remains to be determined. Nonetheless, it appears likely that, under health care reform, the use of various pay-for-performance schemes is likely to increase.

CONCLUSIONS

The safety and quality of care in the United States could be improved substantially. A number of available interventions have been shown to improve the safety of care and should be used more widely; others are undergoing evaluation or soon will be. Quality also could be dramatically better, and the science of quality improvement continues to mature. Implementation of value-based approaches such as accountable care that include pay-for-performance related to safety and quality should make it much easier for organizations to justify investments in improving safety and quality parameters, including health information technology. However, many improvements will also require changing the structure of care—e.g., moving to a more team-oriented approach and ensuring that patients are more involved in their own care. Payment reform focusing on value seems very likely to progress and will likely include both positive incentives and penalties related to safety and quality performance. Measures of safety are still relatively immature and could be made much more robust; it would be particularly useful if organizations had measures they could use in routine operations to assess safety at a reasonable cost, and substantial research is addressing this. Although the quality measures available are more robust than those for safety, they still cover a relatively small proportion of the entire domain of quality, and more measures need to be developed. The public and payers are demanding better information about safety and quality as well as better performance in these areas. The clear implication is that these domains will have to be addressed directly by providers.

FURTHER READING

Bates DW et al: Effect of computerized physician order entry and a team intervention on prevention of serious medication errors. JAMA 280:1311, 1998. [PubMed: 9794308]

Bates DW et al: Two decades since to err is human: An assessment of progress and emerging priorities in patient safety. Health Aff (Millwood) 37:1736, 2018. [PubMed: 30395508]

Berwick DM: Era 3 for medicine and health care. JAMA 315:1329, 2016. [PubMed: 26940610]

Brennan TA et al: Incidence of adverse events and negligence in hospitalized patients. Results of the Harvard Medical Practice Study I. N Engl J Med 324:370, 1991. [PubMed: 1987460]

Chertow GM et al: Guided medication dosing for inpatients with renal insufficiency. JAMA 286:2839, 2001. [PubMed: 11735759]

Institute of Medicine. Report: To err is human: Building a safer health system. 1999. *https://www.nap.edu/resource/9728/To-Err-is-Human-1999--report-brief.pdf.*

Institute of Medicine. Crossing the quality chasm: A new health system for the 21st century. 2001. *https://www.nap.edu/catalog/10027/crossing-the-quality-chasm-a-new-health-system-for-the.*

Landrigan C et al: Effect of reducing interns' work hours on serious medical errors in intensive care units. N Engl J Med 351:1838, 2004. [PubMed: 15509817]

McGlynn EA et al: The quality of health care delivered to adults in the United States. N Engl J Med 348:2635, 2003. [PubMed: 12826639]

Pronovost P et al: An intervention to decrease catheter-related bloodstream infections in the ICU. N Engl J Med 355:2725, 2006. Erratum in: N Engl J Med 356:2660, 2007. [PubMed: 17192537]

Starmer AJ et al: Rates of medical errors and preventable adverse events among hospitalized children following implementation of a resident handoff bundle. JAMA 310:2262, 2013. [PubMed: 24302089]