

What Exactly Is Patient Safety?

Linda Emanuel, MD, PhD; Don Berwick, MD, MPP; James Conway, MS; John Combes, MD; Martin Hatlie, JD; Lucian Leape, MD; James Reason, PhD; Paul Schyve, MD; Charles Vincent, MPhil, PhD; Merrilyn Walton, PhD

Abstract

We articulate an intellectual history and a definition, description, and model of patient safety. We define patient safety as a discipline in the health care professions that applies safety science methods toward the goal of achieving a trustworthy system of health care delivery. We also define patient safety as an attribute of health care systems that minimizes the incidence and impact of adverse events and maximizes recovery from such events. Our description includes: why the field of patient safety exists (the high prevalence of avoidable adverse events); its nature; its essential focus of action (the microsystem); how patient safety works (e.g., high-reliability design, use of safety sciences, methods for causing change, including cultural change); and who its practitioners are (i.e., all health care workers, patients, and advocates). Our simple and overarching model identifies four domains of patient safety (recipients of care, providers, therapeutics, and methods) and the elements that fall within the domains. Eleven of these elements are described in this paper.

Introduction

A defining realization of the 1990s was that, despite all the known power of modern medicine to cure and ameliorate illness, hospitals were not safe places for healing. Instead, they were places fraught with risk of patient harm. One important response to this realization has been the growth of interest in patient safety. It is increasingly clear that patient safety has become a discipline, complete with an integrated body of knowledge and expertise, and that it has the potential to revolutionize health care, perhaps as radically as molecular biology once dramatically increased the therapeutic power in medicine.

Patient safety is now recognized in many countries, with global awareness fostered by the World Health Organization's World Alliance for Patient Safety. And yet there continue to be significant challenges to implementing patient safety policies and practices. One fundamental requirement for adopting any new approach is a clear articulation of its premises and manifestations. Components of patient safety have been expressed by thought leaders, and models have been presented. However, a single rendition that can help a thorough adoption of patient safety throughout health care has not been available. This paper aims to offer that. After introducing salient points in the intellectual history of patient safety, we offer a definition, a description, and finally, a model of patient safety. We call on organizations to adopt a definition and model for patient safety.

Intellectual History of Patient Safety

Critical assumptions in health care were rewritten by patient safety thinking. How to understand why people make errors that lead to adverse events shifted from a single cause, legalistic framework to a systems engineering design framework, and in so doing, it changed forever the way people think about health care delivery.

Limiting Blame

The first quantum leap defined patient safety's entry into health care thought. The realization that adverse events often occur because of system breakdowns, not simply because of individual ineptitude prompted the change. The traditional approach assumed that well-trained, conscientious practitioners do not make errors. Traditional thinking equated error with incompetence and regarded punishment as both appropriate and effective in motivating individuals to be more careful.

The use of this kind of blame had a toxic effect. Practitioners rarely revealed mistakes, and patients and supervisors were frequently kept in the dark. Low reporting made learning from errors nearly impossible, and legal counsel often supported and encouraged this approach in order to minimize the risk of malpractice litigation.¹ This mind-set lent a wary, antagonistic backdrop to the therapeutic interaction.² It also created a locked-in paralysis for all concerned when failure did occur.

Thinking began to change in the 1990s in response to several kinds of new information. First, medical injury was acknowledged as occurring far more often than heretofore realized, with most of these injuries deemed preventable. Second was the idea that "active" errors at the "sharp end"—where practitioners interact with patients or equipment—result from "latent" errors, as demonstrated by James Reason.³ Latent errors are upstream defects in the design of systems, organizations, management, training, and equipment ("blunt end") that lead individuals at the sharp end to make mistakes. To punish individuals for such mistakes seemed to make little sense, since errors are bound to continue until underlying causes are remedied.

Systems Thinking

Thought leaders in health care offered persuasive arguments that errors could be reduced by redesigning systems and processes using human factors principles. These could reduce mistakes through design features, including standardization, simplification, and the use of constraints. One such constraint is a "forcing function," which is a design characteristic that makes error impossible (e.g., incompatible connectors that prevent connecting an anesthetic gas to the oxygen port of an anesthesia machine).

Another corollary quantum leap to view health care as a system took place as people applied engineering design concepts to health care. Some of these systems changes were related to tools and technology, such as using better intravenous pumps or computerizing physician medication prescribing. Others were related to organizations and people, such as training doctors and nurses to work better in teams or including a pharmacist in the team during rounds. Some were more

successful than others, but the important change was that people were thinking of health care delivery in terms of systems.

Interestingly, in earlier phases of medical history, different forms of systems thinking were dominant. However, these forms focused on the biologic systems within the individual patient, rather than on care and interactions between individuals in the environment of care. The notion of humors and the understanding of the circulatory system are two examples from the period prior to the modern scientific era. As the scientific era dawned and the field of medicine began applying the scientific method with success, systems thinking within physiology continued. Perhaps this was helpful, as clinicians took on a systems understanding of the delivery of health care as well.

Initially, perhaps, blunt-end factors were typically thought of as organizational policies and processes that shaped the behavior of individuals at the sharp end-point of service. However, an awareness also emerged of extra-organizational blunt-end factors, including regulators, payers, insurance administrators, economic policymakers, and technology suppliers. These parties often influence and shape incentives and demands within the health care organization. Thus, health care had to be seen as an open, not closed, system, and policy too began to be thought of as a feature of the system.

Transparency and Learning

The idea that adverse events could yield information was not new, but as it was newly applied in health care, it acquired a new potency. The notion that sharing information about medical errors was essential for effective patient safety outcomes became urgent. Commentators asserted that the more error-related information was shared, the better lessons could be implemented industry-wide.⁴ The possibility that knowledge of systems might require an understanding of how things go wrong was demanding attention.

Culture and Professionalism

Clinicians, governing boards, executive leaders, and middle managers of health care delivery organizations were being increasingly encouraged to think in terms of building high-reliability organizations. This required a culture change to one that refrained from assigning “sharp-end” blame for mistakes; that incentivized learning by fully disclosing information about mistakes, failure, and near misses; that trained and provided support to clinicians involved in inherently risky work; and that disclosed all relevant facts to injured parties.^{5, 6}

These transformations in thinking resulted in approaches that were remarkably well-rooted in the essential ethical underpinnings of the profession. The call for safety went directly to the central medical professional imperative to “above all, do no harm.” The value at issue was nonmalfeasance. As a matter of justice, human rights, or the fiduciary obligations intrinsic to the unequal power structure of the provider/patient relationship, the call for systemwide transparency coexisted with fundamental professional standards requiring honesty and disclosure of material facts to the patient.^{7, 8, 9}

Accountability for Delivering Effective, Safe Care

Early Western medical traditions were organized through guilds that kept the special knowledge and skills involved in medical practices a secret.¹⁰ At a time when many medical methods were of dubious foundation, rarely beneficial, and frequently harmful, the challenge of securing the trust of society was significant.¹¹ The primary method was to root out the charlatans. As modern concepts of negligence developed, emphasizing litigation to deter substandard behavior and individual accountability for procedures and actions causally linked to adverse outcomes became embedded in both medicine and law.

In an important parallel development, as treatments became increasingly effective, the medical field began to establish methods for accountability, and the profession's credibility in society rose. The scientific method was essential in that development, and with good reason, medicine has adhered to it. The three-phase approach to establishing the efficacy and safety of new medical therapies—Phase 1, clinical trials to assess safety; Phase 2, clinical trials to ascertain efficacy; and Phase 3, trials to compare it with another standard intervention—was essential, too. The dependence on the randomized clinical trial as the touchstone of the scientific method was critical to that process. The goal was to be sure that medicine was, and was seen as, a clinical research-driven, reliable practice. The effort was successful; society recognized that medicine merited its standing as a profession with specialized expertise to use powerful methods applied appropriately. Consequently, these scientific and clinical research methods and their associated ways of thinking became well entrenched.

The growth of medical sciences also changed standards in medical education, licensure, and peer review. The early apprenticeship model was supplemented by requirements for a phase in which didactically acquired knowledge was transmitted prior to the apprenticeship. As specialties developed, these sought to codify and legitimize their expertise through testing and certification. With the development of safer and more effective surgery, medical care delivery systems began focusing on hospitals; standards for these delivery systems were understood to be necessary. Certification of hospitals and other health care delivery systems followed, often with professional groups, such as the Accreditation Council for Graduate Medical Education (ACGME) and the Joint Commission, serving quasi-government oversight and public protection roles.

The nascent realization that health care, including the clinician and other components, also needed to be accountable for learning from error was harder to grapple with. Faltering moves were made toward tort reform and institutional accountability for safety practices. A model for accountability of clinicians that included accountability for continuous learning set the stage for, but stopped short of, a full rendition of what accountability for understanding and optimally designing safe health care systems required.

Health Care as an Industry

Beginning in the first half of the 20th century, the industrial era phased into the service industry era. Systems thinking was an established part of industrial engineering and applied in production lines and service industries. Yet medicine maintained a separation from these changes. This may have been possible mainly due to medicine's standing as a revered profession with a privileged

relationship to society, but in part, it also may have occurred because both providers and patients protected the one-to-one model of the doctor-patient relationship. Thus, the health care paradigm remained focused on the patient-physician relationship and on a therapy's point of application, rather than on the systems of application. The practitioner was trained and certified to apply therapy at the point of the illness-causing disorder. Even in the more expansive bio-psychosocial model, safety-oriented systems thinking was missing, even though the roles of the patient's immediate relationship circle and of the community and society were acknowledged.

Rising and apparently uncontrollable health care costs, coupled with increasing evidence of poor quality, ushered in the managed care era, along with demands from the public for accountability. Additionally, increased media exposure of preventable medical errors raised troubling questions that propelled a search for new solutions. Leape's earlier publication of the theoretical possibility of applying industrial human-factors engineering concepts to health care,¹² and the subsequent demonstration with Bates and colleagues⁶ of the utility of systems analysis in understanding medication error later that year, provided that new type of thinking. The first conference on patient safety and systems error at the Annenberg Center for Health Sciences in 1996 was a natural next step toward a new type of thinking.

Rethinking Risk

Thought leaders from medicine and policymakers began to carve a new way of understanding risk, new ways to reaffirm relationships with patients, and a new way of addressing the shocking realities that epidemiologic studies, such as Leape's 1994 landmark study, *Error in Medicine*, had presented.¹² A decade earlier, anesthesiology had made substantial improvements by applying systems thinking translated from methods used in aviation and mechanical engineering, but the rest of medicine had failed to generalize it. Quality improvement and risk management had both developed as disciplines within health care, with an emphasis on health services delivery research and measurement. These and other developments produced a readiness for looking at what might be learned and adapted from other high-risk industries and complex organizations.

Emphasizing Teamwork as Well as Dyadic Relationships

Early attempts at systems change revealed one Achilles heel of implementation: dysfunctional relationships between clinicians and other workers. Mirroring some of the developments in aviation—in which a focus on teamwork complemented attention to refinement of mechanical systems—health care began to recognize the importance of team functioning, particularly for communicating across authority gradients. Training in teamwork became a foundational building block for the new field of patient safety.

The discipline of patient safety rejected the concept of health care delivery as an exclusive dominion of the medical profession over the patient-physician relationship. The vision was more inclusive and demanding. It included patient-centered care and the biomedical model, and it focused on interdisciplinary teams and families. It also included the technical and administrative aspects of health care delivery in a complex system.

Defining Patient Safety

As the intellectual history of patient safety developed, it became increasingly important to define patient safety. Thought leaders began to examine their different assumptions. Is patient safety a way of doing things—i.e., a philosophy (with its own explanatory framework, ethical principles, and methods) and a discipline (with a body of expertise)? Or is it an attribute—i.e., a goal and a condition (being safe), a property that emerges from the system? Existing definitions seemed to vary on the question.

Although the Institute of Medicine (IOM) defined safety as “freedom from accidental injury,” patient safety as a discipline or field of inquiry and action has not been fully defined to date in the major consensus statements of the organizations that have propelled its existence. Part of the challenge lies in distinguishing safety from quality, a line that remains important to some, while being dismissed by others as an exercise in semantics. In 1998, the IOM convened the National Roundtable on Health Care Quality, which adopted the following definition of quality that was widely accepted: “Quality of care is the degree to which health care services for individuals and populations increase the likelihood of desired health outcomes and are consistent with current professional knowledge.”¹³

Health care quality problems were classified into three categories: underuse, overuse, and misuse, all of which the evidence shows are common. Misuse was further defined as the preventable complications of treatment. Although the IOM Roundtable was careful to distinguish misuse from error (the latter may or may not cause complications), the misuse category became a common reference point for conceptualizing patient safety as a component of quality.

In 2006, Leape and Berwick observed that, as attention to patient safety has deepened, the lines between the overuse, underuse, and misuse categories have blurred. “It seems logical,” they wrote, “that patients who fail to receive needed treatments, or who are subjected to the risks of unneeded care, are also placed at risk for injury every bit as objectionable as direct harm from a surgical mishap.”¹⁴

The National Patient Safety Foundation identified the key property of safety as emerging from the proper interaction of components of the health care system, thereby leading the way to a defined focus for patient safety, namely systems.¹⁵ Its goal has been defined as: “[t]he avoidance, prevention, and amelioration of adverse outcomes or injuries stemming from the process of care.”¹⁶

Our Definition of Patient Safety

We use the following definition of patient safety:

Patient safety is a discipline in the health care sector that applies safety science methods toward the goal of achieving a trustworthy system of health care delivery. Patient safety is also an attribute of health care systems; it minimizes the incidence and impact of, and maximizes recovery from, adverse events.

This definition acknowledges that patient safety is both a way of doing things and an emergent discipline. It seeks to identify essential features of patient safety.

The Why, What, Where, How, and Who of Patient Safety

Going farther with the definition, each of its components is expanded here to offer a deeper description of patient safety:

Why does the field of patient safety exist? Patient safety as a discipline began in response to evidence that adverse medical events are widespread and preventable, and as noted above, that there is “too much harm.” The goal of the field of patient safety is to minimize adverse events and eliminate preventable harm in health care. Depending on one’s use of the term “harm,” it is possible to aspire to eliminate all harm in health care.

What is the nature of patient safety? Patient safety is a relatively new discipline within the health care professions. Graduate degree programs are currently being introduced in recognition of patient safety as a discipline. It is a subject within health care quality. However, its methods come largely from disciplines outside medicine, particularly from cognitive psychology, human factors engineering, and organizational management science. That, however, is also true of the biomedical sciences that propelled medicine forward to its current extraordinary capacity to cure illnesses. Their methods came from biology, chemistry, physics, and mathematics, among others. Applying safety sciences to health care requires inclusion of experts with new source disciplines, such as engineering, but without any divergence from the goals or inherent nature of the medical profession.

Patient safety is a property that emerges from systems design. Patient safety must be an attribute of the health care system. Patient safety seeks high reliability under conditions of risk. Illness presents the first condition of risk in health care. Patient safety applies to the second condition: the therapeutic intervention. Sometimes the therapeutic risk is audacious, such as when a patient’s heart is lifted, chilled, cut, and sewn during cardiac transplantation surgery. Risk and safety are flip sides of the therapeutic coin.

Patient safety demands design of systems to make risky interventions reliable. Two tenets of complexity theory apply: First, the greater the complexity of the system, the greater is the propensity for chaos. Second, in open, interacting systems, unpredictable events will happen. The better the therapeutic design, the more resilient it is in the face of both predictable and unpredictable possible or impending failures, so they can be prevented or rescue can be achieved.¹⁷ Safety systems include design of materials, procedures, environment, training, and the nature of the culture among people operating in the system.

Berwick and others have collaborated with Amalberti to apply Shewhart’s notion of statistical quality or error levels to health care.¹⁸ Systems are categorized by their level of adverse events. Barriers to progression from one level to another are identified. Interestingly, leaders of high-reliability organizations in other industries view the level of adverse events in medicine as so high that many of them would consider the health industry as existing in a state of chaos. The patient safety discipline seeks systems that can move health care to higher and higher levels of safe care.

Patient safety is a property that is designed for the nature of illness. High-reliability design is a concept that was not originally developed for health care. However, health care has some essential features in common with how high-reliability design has evolved. While often complex and unpredictable, it can have the ultimate high-stakes outcome: preservation of life.

A unique feature of patient care is its highly personal nature. Provision of care almost always requires health care workers to cross significant personal boundaries, both psychological and physical. To protect patient integrity, the health professions have developed codes of professional ethics that guide how best to provide health care without doing dishonor to the ill person. Patient safety designs must allow for these important restrictions, which include confidentiality, physical privacy, and others. At times, these needs conflict directly with the transparency and vigilance needed for optimal patient care, including safety.

Another unique feature is the natural progression of illness. By definition, when illness care begins, something has already gone wrong. Thus, in many medical situations, failure to provide the correct intervention causes harm to the patient. A missed diagnosis of meningococcal meningitis, for example, usually results in patient death. The patient safety discipline acknowledges the need to include harm due to omission of action, as well as the obvious harm due to actions taken.

The vast diversity of possible etiologies and manifestations of illness makes systems design in health care a unique challenge. Nonetheless, the reality is that most conditions are common and of common etiology, which allows for optimal design, if not infallible outcomes. If most patients with a condition such as breast cancer are best treated according to protocol but some require off-protocol, tailored treatment, systems can be designed to meet that need for the majority of protocols with tailoring options.

Patient safety is a property dependent on open learning. Patient safety has another inherent feature that derives directly from its dependence on errors and adverse events as a main source of understanding. It depends on a culture of openness to all relevant perspectives in which those involved in adverse events are treated as partners in learning. In this sense, patient safety espouses continuous cycles of learning, reporting of adverse events or near misses, dissemination of lessons learned, and the establishment of cultures that are trusted to not cast unfair blame. The patient safety field marries principles of adult education and effective behavioral learning with the traditional approaches of the medical profession. Known from its early days as the field that seeks to move “beyond blame” to a culture trusted by all to be just patient safety, patient safety pioneers have pushed for a much deeper understanding of the mechanisms of errors that often lie beyond the actions or control of the individual.

Patient safety advocates turn away from the traditions of the guild in which social standing and privileged knowledge shielded practitioners from accountability. They also reject the defensive posture of old risk management approaches in which physicians and leaders of health care organizations were advised to admit no responsibility and to defend all malpractice claims, whether or not they were justified. Patient safety embraces organizational and personal accountability, but it also recognizes the importance of moving beyond blame in both its

organizational and its personal dimensions, while maintaining accountability and integrity in interactions with patients and families who have suffered avoidable adverse events.

Trustworthiness is essential to the concept of patient safety. The health care system designed for patient safety is trustworthy. This is not because errors will not be made and adverse events will never happen, but because the health care system holds itself accountable to applying safety sciences optimally. Patient safety (as an attribute) prevents avoidable adverse events by paying attention (as a discipline) to systems and interactions, including human interactions, and allowing learning by all parties from near misses and actual adverse events. Through a concerted, conscientious effort, all those involved act to minimize the extent and impact of unavoidable adverse events by creating well-designed systems and well-motivated, informed, conscientious, and vigilant personnel, and by seeking to repair damage honestly and respectfully when it occurs.

Where does patient safety happen? The ultimate locus of patient safety is the microsystem. That is, the immediate environment in which care occurs—the operating room, the emergency department, and so on. It is in the microsystem where the “sharp end” resides, where patient-caregiver interactions occur, where failures of safety emerge, and where patients are harmed. Breaches in safety may have occurred in many blunt-end components, and as described above, events constitute properties of interacting components of the overall system. Therefore, patient safety is irreducibly a matter of systems. Nonetheless, as the setting where the patient receives health care, the microsystem is the locus where the successes or failures of all systems to ensure safety converge.

At the same time, patient safety must be concerned with the entire system. Importantly, patient safety recognizes that the microsystem is inherently unpredictable. Although it takes a mechanistic view of causation, patient safety acknowledges that each microsystem is open in that it can be influenced by another microsystem. This may result in something unpredictable. Thus, for instance, the microsystem of concern in surgical safety might be the operating suite, but if a local emergency demands that two members of the surgical team leave the operating room, the microsystem has been unpredictably affected.

How is patient safety achieved? A number of mechanisms are involved in achieving patient safety, including:

High-reliability design. The fundamental mechanism by which patient safety can be achieved is high-reliability design, which includes many components. Thus, the irreducible unit of patient safety delivery is multifaceted; all components of health care delivery must be integrated into a system that is as reliable as possible under complex conditions.

A unique feature of high-reliability design comes from complexity theory, which notes that open, interacting systems will produce some level of chaos or inherently unpredictable events. High-reliability designs are resilient even when unpredictable events occur.

Additional design features that guide health systems engineers include “lean process” and a notion of breaking through reliability boundaries in leaps from one safety level to another. These

