Caveats for high reliability in healthcare

Peter F. Martelli and Peter E. Rivard
Sawyer Business School, Suffolk University, Boston, Massachusetts, USA, and
Karlene H. Roberts
Haas School of Business, University of California Berkeley, Berkeley, California, USA

Abstract
Purpose – Given the pace of industry change and the rapid diffusion of high reliability organization (HRO) approaches, lags and divergences have arisen between research and practice in healthcare. The purpose of this paper is to explore several of these theory-practice gaps and propose implications for research and practice.
Design/methodology/approach – Classic and cutting-edge HRO literature is applied to analyze two industry trends: delivery system integration, and the confluence of patient-as-consumer and patient-centered care.
Findings – Highly reliable integrated delivery systems will likely function very differently from classic HRO organizations. Both practitioners and researchers should address conditions such as how a system is bounded, how reliable the system should be and how interdependencies are handled. Additionally, systems should evaluate the added uncertainty and variability introduced by enhanced agency on the part of patients/families in decision making and in processes of care.
Research limitations/implications – Dramatic changes in the sociotechnical environment are influencing the coupling and interactivity of system elements in healthcare. Researchers must address the maintenance of reliability across organizations and the migration of decision-making power toward patients and families.
Practical implications – As healthcare systems integrate, managers attempting to apply HRO principles must recognize how these systems present new and different reliability-related challenges and opportunities.
Originality/value – This paper provides a starting point for the advancement of research and practice in high-reliability healthcare by providing an in-depth exploration of the implications of two major industry trends.
Keywords Safety, Quality, Healthcare, High-reliability organization (HRO), Increasing relative ignorance, Reliability seeking
Paper type General review

1. Introduction
Efforts at high reliability organizing are diffusing with increasing speed among healthcare delivery organizations (Chassin and Loeb, 2013). Given the variety of approaches to “doing high reliability” in healthcare, evidence on these practices is only beginning to accumulate, and research may seriously lag practice (Vogus and Iacobucci, 2016; Vogus and Singer, 2016). Likewise, broader high reliability theorizing and practice may have diverged. This paper identifies some apparent theory-practice gaps concerning high reliability in healthcare, and articulates implications for both research and practice.

Early in its conceptualization, a seminal paper argued that high reliability might be “working in practice but not in theory” (LaPorte and Consolini, 1991). The paper highlighted researchers’ surprise at low failure rates in certain organizations whose components were characterized by high interactive complexity and tight coupling. It concluded that, notwithstanding these real-world successes, contemporary theories to explain the phenomena remained inadequate. The obvious implication was to theorize more about how these organizations function, particularly regarding “behavioral responses in decision making” (LaPorte and Consolini, 1991).
making, in patterns of authority, in processes of discovery [...] and in responses to regulation” (LaPorte, 1994). This ensuing work noted, somewhat pessimistically, that “we know less of what we need to know, even as we know more than we did” (LaPorte, 1994). In other words, the field is subject to “increasing relative ignorance,” where theorizing lags empirical change – not only in terms of vast technological advancement, but also through changes imposed by the socio-economic environment (e.g. new regulatory schemes) and changes inflicted by principals upon agents (e.g. the growth of patient-as-consumer). Therefore, it is possible that we know more than we did in that early phase, yet still less than we need to in order to satisfactorily analyze, explain and predict.

The warning of increasing relative ignorance is particularly apt in healthcare. This stems in part from the fact that comparatively little of the foundational high reliability organization (HRO) research has been in healthcare. And even as it undergoes the changes noted above, healthcare also differs in consequential ways from other industries viewed through the HRO lens. For example, in part because every patient is unique and presents uniquely, healthcare’s core processes are subject to greater uncertainty and variability. While a stream of integrative and implementable research on high reliability in healthcare is emerging, important questions have yet to be sufficiently addressed. In what ways have the conditions of healthcare management practice changed faster than theory can describe? To what extent is high-reliability theory reifying despite changes to constraints and parameters of interest at the “sharp end” of work?

LaPorte’s prescient analysis noted above seems to apply to our current state in healthcare: the socio-economic environment is moving at a blistering pace and principals (i.e. patients) are asserting themselves in unprecedented ways. Using this as a framework, the following essay identifies and characterizes key instances of these changes and discusses the potential consequences for our grasp of reliability in this setting. The argument is intended to provoke clearer consideration of the caveats on high-reliability healthcare, especially with respect to the boundaries of the system to be managed and the scope and extent of the outcomes subject to reliability management. Two sections on implications for research and practice draw out these questions further and present thoughts for future attention.

2. Conceptual framework
High-reliability organizations (HROs) spark scholarly interest because “there are neither theoretical nor empirical reasons to expect failure-free operations in large organizations that attempt to perform demanding tasks with little margin for error” (LaPorte, 1996). HROs are “defined by their unique ability to operate high-hazard technological systems in a nearly error-free manner” (Roberts, 1990; Vogus and Welbourne, 2003). In HROs, “error avoidance and safety are as much a part of the bottom line as is productivity” (Desai et al., 2012).

A systems perspective is central to understanding the HRO, which is conceptualized as a complex system with high interdependence among its parts. A small change, anomaly, error or failure of inputs or components can trigger a chain of events leading to catastrophic system failure. Given the foundational role of systems in HRO research, the applicability of research findings depends in part on clearly and appropriately delineating the boundaries of the focal system. The objects of early HRO research were relatively closed systems, e.g., aircraft carriers and power plants. Healthcare increasingly offers a contrasting picture: an episode of patient care often occurs in a more open and fluid microsystem or “healthcare system” spanning various organizational boundaries.

Even within its agreed-upon boundaries, the organization is itself a system of systems. To that end, HRO research also must be clear about “reliability with respect to what?” What is the system that is to be made reliable? In early HRO research on aircraft carriers, for example, reliability was achieved with respect to the human and mechanical flight deck...
systems supporting aircraft takeoff and landing. Analogously, in healthcare, what is the system that supports favorable surgical outcomes or safe medication administration?

In addition to the need to clarify the “organization” part of “highly reliable organization” is the need to clarify what “highly reliable” means. In HRO, reliability is ultimately measured in relation to failure rates. Yet even early on, Rochlin (1996) made certain to note that the term high is “a judgmental and not an absolute variable” – because what constitutes high reliability is context-specific and socially determined, it can mean zero failures or a very low rate of failures. Schulman (2002, 2004, 2014) distinguishes “precluded-events” reliability from “marginal” reliability. In the precluded-events reliability frame, even one failure is catastrophic; in the marginal reliability frame, zero failures may be the aim, but a non-zero failure rate is tolerated. Failures in healthcare typically stem from care delivery failing to achieve intended results, i.e., harm to the patient or materially less benefit than intended. As with other industries, threshold failure levels for high reliability are context-specific and subject to debate. For example, when does the pursuit of high reliability mean precluding “never events” such as wrong-site surgery, and when does it mean reducing undesirable events (e.g. superficial surgical site infections) to very low rates? While “zero harm” is the ideal, it is not currently achievable or even desirable (at least from a utilitarian perspective) in all healthcare delivery microsystems.

The reliability of an organization can be assessed in terms of outcomes or in terms of the organizational structures and processes that are predictive of those outcomes. In healthcare, relevant outcome measures include adverse events (patient harm) such as postoperative infections or adverse drug events (ADEs), and indicators of treatment failures such as hospital readmissions. An additional challenge to assessment of the reliability of healthcare organizations is their nature as collections of units with varying degrees of autonomy. Reliability therefore is not monolithic: the “problem is not that health care is not reliable or resilient at all, but that huge variability exists within teams, within organizations, and across the system. The hospital that contains centers of excellence may have other units in which outcomes are poor or even dangerous” (Vincent et al., 2010).

The management implications of this distinction are considerable. Managing for precluded events is characterized by non-probabilistic (every-last-case) management and non-fungible resources to prevent that defined failure (Schulman, 2014). In contrast, marginal reliability is a continuous, rather than categorical, variable characterized by probabilistic (run-of-cases) management and by fungible resources which are leveraged in tradeoffs across competing reliability priorities (Schulman, 2014). HRO is more accurately characterized dynamically, as a process, than statically – i.e., as organizing rather than organization. As Klein et al. (1995) have noted: “No claim is made either by the organizations or by those studying them that they are immune to catastrophic accidents. In fact, the reason their managers devote so much effort to avoiding accidents is their inherent vulnerability to such accidents.”

The following sections explore recent changes in healthcare policy and practice in order to highlight several unanswered questions about high reliability in healthcare. The analysis is structured according to the pressures identified by LaPorte (1994), presenting evidence of trends affecting reliability in this setting and describing the “relative ignorance” faced as a consequence. Each section concludes with promising potential routes toward knowing “more of what we need to know” in order to make reliability a reality in healthcare.

3. Changes imposed by the socio-economic environment: uncertainty, integration and interorganizational reliability

3.1 Accountable care organizations (ACOs) as drivers of integration

Growing understanding of the failures of fragmented healthcare services has led to a push for more integrated delivery systems. Among the vehicles for integration in the USA,
perhaps the most visible and clearly articulated is the ACO. ACOs have quickly achieved broad geographic and market penetration. Over 70 percent of the 28m plus ACO members are in Medicare/Medicaid. This market dominance often leads medicare ACO program changes to diffuse to commercial programs (Gandhi and Weil, 2012). ACO-related regulatory change, stemming in part from the 2010 Affordable Care Act, is rewriting the institutional environment and driving behavior change through economic incentives, legal obligations, regulatory sanction and cultural stimulation. Patterns of market entry and other trends suggest that ACOs will continue to form (Frech et al., 2015; Slonim and Maraccini, 2016). A shared understanding is emerging of what an ACO is and what it can and cannot achieve, offering system actors a roadmap for coordinating strategies. However, volatility and uncertainty in the industry and in the federal and state policy context leave “the rules of the game” unsettled. Shifts toward integration are likely to be piecemeal as organizations keep one foot in current ways of doing business (Gold, 2013). Strategic decision-makers may not embrace or fund ACO-type integration sufficiently for it to be effective, efficient or reliable.

The ACO model entails “a governance structure and processes of accountability [for both cost and quality] where providers […] operate collaboratively within the overarching ACO boundary” (Addicott and Shortell, 2014). The ACO is designed as an organizational integrator accepting responsibility for the “Triple Aim” of population health, experience of care and lower per capita costs for an assigned population (Berwick et al., 2008). Compared to earlier integration efforts, such as health maintenance organizations, ACO models assume that recent developments in data collection, measurement and analysis will afford today’s ACOs and their payers greater capabilities to assess, incentivize and improve performance on the Triple Aim.

Under typical ACO payment models, total costs for assigned patients are fully or partially liable to the mutually accountable providers. The theoretical crux is that shared accountability promotes greater coordination and integration of structures and processes across the care continuum as well as investment in population health. Because costs are borne against an expected per-member payment to the ACO, it becomes irrational merely to shift costs to other parties within the ACO. Evidence supports the importance of provider integration to ACOs: ACO market entry has been higher where provider integration was already higher (Chukmaitov et al., 2017) and successful alliances among providers of both clinical and social services have proven key to ACO success to date (Shortell et al., 2015). Evidence on ACO performance is mixed, however (Blumenthal et al., 2015; Shortell et al., 2015). One study found “few examples of physicians being held to account as a collective and therefore only limited evidence of collaborative accountability impacting on behavior change” and the authors suggest that an optimal integrated governance arrangement has yet to emerge (Addicott and Shortell, 2014). Thus, in addition to general instability in the policy environment, the ACO form’s newness and lingering design ambiguities pose at least temporary barriers to full investment.

3.2 Interorganizational reliability
This environment has many implications for reliability. One key question for researchers and practitioners is: where does (or should) reliability reside? This was relatively straightforward in early HROs, which approximated the closed rational system archetype (LaPorte and Consolini, 1991). Decisions were made, problems solved and resources consumed with little need to reference an environment beyond the immediate organizational system. The organizational boundary was relatively clear, and reliability “resided” in the internal capacity to avert precluded events. Many organizations currently aspiring to high reliability lack the relative luxury enjoyed by these earlier “ships in a bottle”; their processes occur in open systems that disregard traditional organizational boundaries.
Examining where reliability resides entails examining system boundaries. Managing healthcare delivery for reliability will mean managing care interorganizationally.

There are two fundamental ways to gain traction on interorganizational reliability: first, to reassess the nature of reliability under this constraint; or, second, to redefine the nature of the system boundaries.

First, Roe and Schulman discriminate “classic” HRO as a particular type of organization from the broader category of high-reliability management (HRM) organizations, which fall along a continuum with HRO at one extreme (Roe and Schulman, 2008; Schulman, 2017). The character of problems, input and output variance and formal vs tacit knowledge requirements all differ between HRO and other HRM organizations; hence, the nature of reliability differs as well. HRM, they contend, “is the provision of what society considers to be critical services safely and continuously, even during peak demand and turbulent times, and often under conditions of increasing interconnectivity” (Schulman and Roe, 2014). The interconnectivity aspect is a key contribution of HRM theorizing. Researchers and practitioners may benefit from viewing healthcare integration through an HRM lens. Citing the US Department of Homeland Security as an example, Roe and Schulman (2008) point out that the “Health Care and Public Health Sector” is 1 of the 16 critical infrastructures requiring management. There may be particular and significant challenges to managing reliability in this network context, where organizations are subject to “institutional fragmentation, automation, and liberalization processes” (van Eeten et al., 2010). When interorganizational integration is poor, the requisite information and resources to construct long-term, anticipatory plans for reliability are inadequate, so in an HRM context, operational requirements shift to compiling these cognitive and material resources in order to manage “just in time” (Roe and Schulman, 2008).

Second, integration might be explored by viewing healthcare organizations as interdependent subunits within a larger, even societal, system striving to operate reliably. Scholars concerned with this have recognized that HROs are a special case of what have been termed large technical systems, in which the “most significant properties […] are increasingly intensive knowledge requirements, tightening patterns of functional interdependence within major productive or service segments and expanding networks of cooperation and control” (LaPorte, 1996). Recent research suggests that “reliability increasingly is not a property of single organizations as researched in the past,” as many organizations exist in “a world where infrastructures and supply chains are interconnected and networked together” (Schulman and Roe, 2014). The Center for Catastrophic Risk Management’s ongoing study of Interdependent Critical Infrastructure Systems supports this assertion (Schulman and Roe, 2016; Roe et al., 2016). As interconnectedness increases, specifying organizational boundaries becomes more difficult. Examining reliability requires addressing cross-organizational interfaces that provide and secure resources, including knowledge, subject to a range of governance arrangements. In other words, “control operators in each infrastructure have to depend upon other control operators in real time” (Schulman and Roe, 2016; Roe et al., 2016) such that reliability “resides” across a system of organizations. The great variety of services delivered by healthcare organizations complicates the task of building reliability interorganizationally. The “reliable with respect to what” question will have multiple answers (e.g. emergency department services vs joint replacement vs an episode of congestive heart failure), each service might be delivered by a differently bounded system, and decision-makers will set different reliability targets across the services (precluded events vs marginal reliability).

Subsystems are likely to be, paradoxically, both autonomous and interdependent. While autonomous in many respects, “the outcomes they reach are often determined by the nature and distribution of resources among them as well as the way in which these resources are used” (Grabowski and Roberts, 1997). For example, researchers are seeking to learn how
“reliability enhancing sub units [are] interdependent with their larger organizations […]. [and how] organizations that can fail disastrously [are] linked to other organizations that can either help them fail, or fail alongside” (Roberts, 2012). Rochlin (2012) identifies the importance of considering “whatever contribution safer or more reliable operation of our sub-organizations to the nested series, or the overall whole” might make.

One analysis suggests that if vigilance, shared cognition and constrained improvisation are the general challenges of reliability, then HRO implementation in interorganizational systems will require multiple facets of organization. Those facets include: cultural compatibility by virtue of “cultural consideration in partner selection, long-term relationships, and development of trust”; rich information sharing, including accident and near-miss reporting; an appreciation that contractual relationships employing financial incentives to promote safety may be “counterproductive”; and joint training in order to “develop shared cognition, and establish parameters for constrained improvisation” (Madsen, 2009). Additionally, we might gain some traction by examining the implications of an exchange theory or resource dependence approach, following LaPorte and Consolini’s (1991) observation in “traditional” HROs that “both within and among critical operating groups [seen in HROs], patterns of high degrees of interdependence, i.e., the mutual exchange of social and financial resources, appears particularly evident.”

3.3 Governance, central control and subunit autonomy
Organizational governance is another facet of reliability in networked organizations: echoing Grabowski and Roberts’ (1997) observation on autonomy and interdependence, recent research on interorganizational networks has identified features that appear to support high reliability, including dynamic network governance that switches between democratic, multi-layered control and temporary, “assertive” central control (Berthod et al., 2017). An essential element of HRO is the subunit’s awareness of, and responsiveness to, new hazards and changing local operating conditions. The organization’s standard procedures may be set aside in order to maintain reliable operations. Yet, if the standard procedures themselves are not updated and communicated in the broader organization when new practices take hold, this reliability-enhancing local adaptation may become a reliability-reducing “drift” of local practice away from the shared understanding that sustains reliability across subunits (Snook, 2000). Just as newly integrating organizations have an up-front need for compatible culture, information sharing, appropriate economic incentives and joint training, they also may need a dynamic integrative function that monitors fluctuations and discrepancies in practices – i.e., “drift” – across subunits (Pettersen and Schulman, 2016) and updates standard procedures as appropriate.

Interdependencies among contractors, suppliers/vendors, unions, regulators, customers and competitors increasingly cross organizational boundaries (Madsen, 2009). However, unlike many other products, achieving the product of “good health” is almost always a reciprocal task, requiring mutual adjustment between interdependent parties (Thompson, 1967). Those responsible for providing “good health” constantly make customized, task-related choices that constrain each other’s options and affect outcomes. Safety culture, teamwork and relational coordination – all promoted in healthcare in recent years – represent ways to accomplish tasks where plans or standards are not enough, and may need to be built across organizations. Another route is assigning a special boundary-spanning, reliability-oriented role to an individual who attempts to ensure that all responsible parties are interfacing with the patient and each other at the right times. This role approximates the “reliability professional” (Schulman, 2004; Roe and Schulman, 2008), and has existed in healthcare in various forms and names, notably as a disease manager for chronic diseases (Nash and Todd, 1997) or a care coordinator more generally (McDonald et al., 2007), as patient navigators to assist in appointment
scheduling and conveying information (Parker and Lemak, 2011), and perhaps even, if patient financial health is also considered, as a “patient advocate” (Patient Advocate Foundation, 2016).

The negative consequences of interorganizational integration might include failures either of subgroup optimization or of centralized coordination. Centralized coordination can fail when resources or authority invested in the coordinator are inadequate, or there is an untested assumption that assigning responsibility to a coordinator will bring improvement. For instance, frontline staff may become disillusioned when safety issues they raise are either misunderstood (Singer et al., 2013) or not followed up (Rotteau et al., 2014). Similarly, a partially tongue-in-cheek “No Reporting” campaign was proposed for Swedish aviation, stressing the failures of centralized event-reporting systems, under the banner, “we will report again when you have acted on the 70,000 reports that are already in the system” (Cook, Personal communication). It could also be that failures to fully align values or incentives across an integrated system lead to the economically rational maximization of discrete unit goals agnostic to the system’s desired outcomes. An organization may experience decidedly variable reliability across subgroups, and “unfortunately, the perceived reliability of a large and diverse organizations [sic] will usually depend on the worst performing visible sub-unit” (Rochlin, 2012). While society’s perceptions remain critical for large technical systems, similar perceptual error could easily extend to managers controlling resource allocation. For instance, attention to the missed performance benchmarks of individual departments could, counterintuitively, redirect resources away from system reliability in favor of ensuring the “success” of the visible subunit. This suboptimization “whack-a-mole” not only redirects vital resources, but also implicitly – and, in a dynamic system, erroneously – assumes a static state of system reliability, where the ultimate role of management would be simply to prevent deviations within functional subunits. Moreover, system reliability can only be perceived to the extent that poor performance is visible, or even calculable, given our far-from-complete understanding of disease and its treatment. Similarly, Roberts (2012) states that constant attention to the worst performing visible subunit without high-level system intervention could lead to “errors of the third kind” (Bea et al., 2009), wherein managers focus on precisely solving an improperly specified problem.

3.4 Reliability competes with other strategic priorities

Another challenge of central coordination is differentiating the risks deemed by the boardroom to be necessary for success from the risks of failure to be avoided by the control room. In other words, to what extent can high reliability be central to business strategy when market positioning and financial performance are critical? The disconnect could go either way: greater risk-taking in service of goals other than reliability could occur either in the central office or at the subsystem level. Swedish Health Services in Seattle provides an unfortunate case in point. The system “designated patient safety as its No. 1 priority [...] [by committing] to a cultural transformation grounded in comprehensive high-reliability training for physicians and staff” (Birk, 2015). However, surgeons were “incentivized to pursue a high-volume approach with contracts that compensate them for large patient numbers and complicated surgical techniques” (Ehley, 2017). This “created a perception that what really matters at Swedish is vast RVU [relative value unit] production without concern for the means by which it is achieved,” eventually leading to serious patient safety issues and a public relations crisis (Baker and Mayo, 2017).

In addition to the potential for reliability, the integration sought by ACOs serves multiple aims, including lower transaction costs and insurance costs and reduced exposure to regulatory cost or sanction (Mick and Shay, 2016; Schulman, 2002). While the mechanisms driving financial integration may likewise drive clinical integration, it could be as Gaba (2000)
presaged, that “[...] such consolidation will focus strictly on cost-cutting or will generate increased organizational complexity that can actually make errors more likely.” For those who plan and manage integrated healthcare systems, the human and other resources essential to a highly reliable system may not appear to be “value-adding” at budget time. In retrospect, Schulman (2017) comments that:

[...] it is no surprise that most of the HROs observed [in the original studies] were either public agencies or functioned in highly regulated markets. Regulation in many cases stabilized their environments and protected these organizations from strong market competition that would have put pressure on them to cut reliability costs while at the same time constraining any competitors to the same regulatory requirements.

4. Changes inflicted by principals upon agents: the patient-as-consumer alters system reliability

Two relatively recent developments are changing the patient’s role in the reliable delivery of healthcare services: a social and economic shift toward “patient as consumer” and a policy shift toward “patient-centered care.” In most well-developed healthcare systems, the patient has explicitly not been an active economic player. Public and private insurance have spread the high and difficult-to-predict financial risks of disease and enabled access to services. As a risk-shifting intermediary, health insurance has generally ensured that the patient has neither seen the actual price nor borne many of the true costs of care. Moreover, the healthcare “product” is unlike other products. In particular, individuals often lack sufficient knowledge to select the specific services and goods that optimally meet their health needs. Patients therefore enter into agency relationships with providers and payers, through which most purchasing decision making is transferred from patient to provider/payer. These agency relationships make the patient a special sort of consumer – one who rarely has the combination of opportunity, incentive and capability to choose between options based on price or quality, and who bears little to no financial risk for what is consumed. In essence, the patient has not really been a consumer at all. Therefore, from a narrow market perspective, there should be no expectation that goods are allocated efficiently or that patients receive the care they need or desire from the marketplace.

In order to bring rationality to the system, market players are investing in transforming the patient into a real consumer. For example, payers are experimenting with payment designs, such as consumer-directed health plans or value-based insurance design, in order to expose patients to greater economic risk (i.e. bear more of the costs of care) and make them more sensitive to the economic consequences of purchasing choices, or to “nudge” them into choices that maximize the benefits derived. These and other models aim to empower consumers inasmuch as consumers represent a mechanism to drive prices down and/or quality up in a market context. To the extent that patients can judge quality and cost, they can make informed choices – which research suggests tend to favor high value care (Hibbard et al., 2012).

However, empowering patients to be “consumers” is also relevant because it promotes certain values that many consider important from a quality-of-care perspective in addition to an economic perspective. Based on discussions leading up to Crossing the Quality Chasm, the Institute of Medicine (2001) defined quality as “the degree to which health services for individuals and populations increase the likelihood of desired health outcomes and are consistent with current professional knowledge.” Of the several provocative aspects of this definition, the idea that a patient’s “desired” outcomes should play a major role in clinical decision making was perhaps the most troublesome to the medical community at the time. The physician as an autonomous professional and agent of the patient, responsible for purchasing and delivering healthcare, had rarely before been challenged with the true...
“patient-centeredness” demanded by this definition (Charles et al., 1999; Berwick, 2009). Moving to a patient-centered model requires inviting the patient to express individual preferences and to expect those preferences will be taken into consideration: “nothing about me without me” (Institute of Medicine, 2001; National Patient Safety Foundation, 2003). Thus, the growth of consumerism is aligned with a growing emphasis on patient-centeredness.

Aspects of patient-centeredness include showing respect to patients, sharing information and cultivating both the participation and collaboration of patients and families (Institute of Medicine, 2011). Principally, it consists in the practice of shared decision making, whereby two experts are deciding care – namely, the provider who understands the medical evidence and the patient (and family) who understand their values and preferences (Charles et al., 1997). Together with increased access to medical information on the internet, patient support groups and the rise of wearable technologies that allow individuals to monitor basic health indicators, the devolution of clinical decision making to meet the patient’s desired medical outcomes is changing the nature of the medical transaction. In turn, it is also changing the nature of the system of care and the responsibility for safety and quality.

To understand why this might pose a challenge for healthcare reliability, consider the question: reliability with respect to what? “What’s the system being managed so as to prevent system failure?” (Schulman and Roe, 2014). Several issues might warrant further inquiry.

First, consider reliability with respect to the individual patient, who is the “system of interest” (Schulman, 2014). This traditional medical conceptualization of reliability stems from the tradition of a single patient interacting with a single physician. Patients – both their bodies and behaviors – are far more variable and less predictable than the mostly standardized inputs to classic HROs. Moreover, patients are not mere “objects” in a healthcare production line, and even less so as empowered consumers. When shared decision making is taken seriously, patients take on a dual role as both “objects/subjects of the activity,” whereby they are simultaneously “both the one who has to participate in taking decision, and the one to which the outcomes of that decision must be applied” (Bagnara et al., 2010). In addition to presenting clinical surprises, the patient also intervenes in the decision process, “bringing in its own variability” (Bagnara et al., 2010). This variability often results from patients’ interactions with a complex system that they are ill-equipped to comprehend, and can lead to the patient being a cause of both “action errors” (i.e. incompatible patient behaviors) and “mental errors” (i.e. failures in patient thought processes) (PSNet: The Agency for Healthcare Research and Quality’s Patient Safety Network, 2017). Given that the levels of both understanding and engagement by patient/family can be difficult to predict across cases or even within cases over time, expecting patients to accept responsibility for difficult, emotionally charged and highly consequential decisions could be viewed as an imprudent model for a robust safety program (PSNet: The Agency for Healthcare Research and Quality’s Patient Safety Network, 2017; Lyons, 2007).

Customizing care delivery to the patient, a cornerstone of patient-centeredness and of patient-as-consumer, creates another source of variability. If providers can implement niche strategies around subpopulations with similar desired outcomes, this variability might be desirable both for patients and for practices. However, customization of a given service “ideographically, patient by patient” may also introduce challenges both in producing and in measuring reliability from the perspective of whether healthcare procedures “actually achieve low but stable output variance across a run of cases” (Schulman, 2004).

Second, the system to be managed in the face of consumerism and patient-centeredness may be the provider organization. Implicit in organizational theory is that the first goal of all organizations is survival. A saying in healthcare is, “no margin, no mission.” Variability in needs, preferences, actions and other inputs from the patient can put the patient’s interests
at odds with those of the organization. One provider survival tactic has been to shift the costs of that variability to patients. Just as the patient-as-consumer historically played a weak role in markets for healthcare services, the patient was largely and perhaps unwittingly “written out of the system” by healthcare professionals seeking to produce high volumes of good quality, billable units of service such as office visits and hospital stays. The consequences of patient variability were externalized from provider to patient. For example, if the standard treatment, appropriate given the patient’s principal diagnosis and comorbidities, was not appropriate in the context of a particular patient’s life, then the cost of the suboptimal clinical outcome was likely externalized to the patient, perhaps chalked up to “patient noncompliance.” This sort of “provider maximization” at patients’ expense also looks like poor service that a consumer would not tolerate. Reimbursement models are shifting toward rewarding “value,” with revenue driven less by service volume and more by the extent to which services achieve outcomes including quality of patient experience. Consumerism and patient-centeredness will likely force professionals and organizations to reckon more with the variability that exists both across patients and within individual patients over time, rather than push that responsibility back onto patients.

Third, patients-as-consumers may not be equipped to play their designed roles in managing larger systems for reliability. In a study of accountable care, Addicott and Shortell (2014) note that “recognition appears to be growing that to control costs while maintaining or improving quality will require greater emphasis on patient activation and engagement in their care and in the design of the delivery system itself.” Yet they “saw little evidence of patient involvement in governance forums and accountability mechanisms, beyond, in the latter case, traditional patient satisfaction surveys.” Furthermore, no matter how activated or engaged patients are, to the extent that the system attempts to drive that behavior by placing greater financial risk on patients, then an unintended consequence is likely to be greater affordability barriers to care and prevalence of medical bankruptcy (Collins et al., 2015). To be healthy but destitute is hardly a good outcome of care. Paradoxically, efforts to include patients may exclude them. In sum, while recent reports bear out that providers increasingly view patients as colleagues in healthcare delivery and incorporate their advice in decision making (McCluskey, 2015; Luthra, 2015), given that health is unlike other products and healthcare is very far from a perfect market, it is unclear whether the patient can ever become a real consumer.

5. Implications for research

HRO research began with the question, “How do certain complex, tightly coupled organizations achieve failure rates that are much lower than organizational theory would predict?” Given a lack of healthcare industry consensus on HRO benchmarks (i.e. what constitutes low failure rates and by which measures?), assessing the prevalence of healthcare HROs remains difficult. Nonetheless, significant resources and effort are being devoted to healthcare HRO. There may be increasing relative ignorance between HRO theory and its practice in healthcare, and this paper suggests that theorists address this gap by continuing to ask that original question in the context of rapidly evolving healthcare delivery organizations. Researchers might seize this as an opportunity to rethink reliability concepts as well as explore them in novel organizational settings. Two areas of change have been the focus of this paper: attempts to integrate formerly fragmented organizational entities; and engagement and empowerment of the patient-as-consumer.

Integration of healthcare delivery units raises questions about interorganizational reliability. What is the nature of reliability in such a system? Is classic, precluded-events reliability achievable? Do highly reliable systems employ HRM methods, focusing on maximizing reliability across multiple intersecting subsystems and on just-in-time responses to threats to reliability? How does failure unfold within and across local subunits and

Caveats for high reliability in healthcare
centralized control/coordination units? What is the interplay of autonomy and interdependence among sub-units? How are reliability-enhancing information, human and material resources distributed among sub-units? How are interorganizational communication, coordination and compatibility achieved? Do reliability professionals play a role? How are competing demands for throughput/profitability and reliability handled?

The patient-as-consumer also raises important questions for research. How does the answer to “reliability with respect to what” change when the patient’s voice is stronger in the determination of desirable outcomes? How do the challenges of building and sustaining reliability change when the patient’s decisions and actions are more central to the process, and when the system can no longer offload responsibility for dealing with the patient as a source of variability and a potential locus of unreliability?

On a wider scale, both of these pressures are indicative that a new institutional era in healthcare may be arising. Scott et al. (2000) argue that since the Second World War, the USA has experienced three institutional eras shaping the legitimate interpretation of problems and solutions in healthcare: the era of “professional dominance” of providers from 1945 to 1965, when the dominant logic was medical care quality; “federal involvement” from 1966 to 1982, when the dominant logic was equity of access; and “managerial control and market mechanisms” beginning in 1983, when the dominant logic has been efficiency of service provision. Given the arguments above, researchers would do well to consider whether we are entering an era of “market integration” in which healthcare providers come together to deliver coordinated care, and the dominant logic is one of patient and consumer empowerment. If this is the case, then it will be especially important to examine where reliability resides and how to ensure it under these new constraints – and even to rethink how “public” healthcare organizations are, to whom they are accountable, where the system and policy boundaries lie, and what it means to operate in this industry.

6. Implications for practice
Trends toward system integration, accountable care and patient-centered care are evident in many countries (McClellan et al., 2017; Pollock and Roderick, 2018; Wang et al., 2014; World Health Organization, 2017). While regulatory context, degree of state control, financing mechanisms and other factors vary across countries, within the functional areas of healthcare delivery organizations and systems, there is considerable consistency. As evolving socio-economic and technological contexts motivate efforts at greater clinical integration, practitioners concerned with reliability should attend to the challenges and opportunities associated with interorganizational reliability. First, is the question: reliability with respect to what? It is important to be both clear about which processes are to be made more reliable and cognizant of the bounds of the systems where these processes reside. A considerable amount of classic HRO knowledge and practice derive from relatively closed organizational systems; the integrated delivery systems envisioned by designers of today’s ACOs look very different from those earlier organizations.

Second, interorganizational reliability likely demands more than the types of integrating features that support shared financial risk across entities. It will be important to ask whether mechanisms of financial integration in fact support or undermine reliability. Reliability must be built across, not just within, groups and organizations. What is the interface when patients and their medical information flow across the system? How are gaps and hazards identified, communicated and responded to, interorganizationally? How is the problem of “reliability drift” – locally adaptive practices diverging from knowledge shared across organizational boundaries – monitored and resolved (Pettersen and Schulman, 2016)? What is the right balance between local autonomy and central control and how is this managed in a dynamic way? How do the multiple entities in the system develop cultural compatibility, rich information sharing, appropriate economic incentives, functional
governance and functional organizational politics? What are the roles of reliability professionals in the system?

Third, how reliable must a given process or subsystem be? Is it a matter of precluded, “never” events, or a matter of keeping failure rates to an acceptable level? With ever-expanding bounds on the system that is to be made more reliable, for parts of the system there may be a corresponding shift from precluded-events reliability to avoided events or even less stringent reliability as well. A large, complex health system could not amass the resources to achieve zero errors of all types in all parts of the organization. There will be different tolerances for error, depending on the error’s type and location. For example, achieving zero deaths attributable to adverse drug events (ADEs) requires a different level and type of organizational vigilance, and different metrics and resources, compared to minimizing a broader and less harmful category of ADEs. Awareness of the “how reliable” question also could fuel discussion about managing tradeoffs across system partners including patients.

Fourth, as the industry simultaneously moves toward greater patient engagement and patient-centeredness, how do organizations apply knowledge about high reliability to this changing principal-agent relationship? How will responsiveness to patient inputs to clinical decision making – i.e., patient preferences and patient-provided information – be balanced with prudent awareness of the limits to the validity and reliability of those inputs?

7. Conclusion

This paper has proposed some areas where there may be “increasing relative ignorance” between HRO theory and current challenges to practice in healthcare. The concept of HRO taking root in healthcare appears to be deeply associated with aviation and nuclear power (Martelli, 2018). This association has led to wide diffusion and institutionalization of an approach that was originally appropriate to solve a problem within a highly localized hospital setting, e.g., an operating room or emergency department. An increasing amount of healthcare – and the associated hazards – occurs outside these settings and across multiple units and organizations, in settings as diverse as community health centers, convenient care clinics, physical therapy offices, nutritionist kitchens, long-term care facilities or the patient’s home. It is highly likely that theorizing of high reliability is not matching the speed of change in practice.

It is possible that the healthcare delivery industry’s interest in HRO is a managerial fashion that will fall as quickly as it has risen. Alternatively, if HRO efforts begin to achieve their aims, they may become budgetary victims of their own success. LaPorte (1996) observes, “One of the curious things about HRO operations in our society is that when either the consensus about their value declines or economic resources in general become more dear, reliability regimes are more difficult to sustain, especially after conspicuous success and/or as system resources become relatively more scarce.” However, if high reliability is to be the “next stop” on the journey to patient safety, it is important to acknowledge that “[i]n health care, high reliability has come to be used as a proxy for safety, although reliability and safety are not truly equivalent” (Sutcliffe, 2011). This problematic conflation arises not only from indifference to scope conditions, but also from the technical, economic and institutional forces that generated the perceived need to “become” an HRO. “True” HRO has proven tremendously difficult and typically very costly. In healthcare, we should be mindful of “the intrinsic costs and difficulties of seeking to continuously to achieve failure-free performance in large organizations and the theoretical impossibility of assuring it under all conditions” (LaPorte, 1996). Reconsidering high-reliability healthcare – for example, in terms of HRM practices and marginal reliability – might yield an opportunity to bring theory and practice closer together under a reliability-seeking orientation and in turn hopefully stem normative adopters from looking for turn-key solutions and certifications.
References


Gandhi, N. and Weil, R. (2012), “The ACO surprise: there’s a common belief that Medicare’s ACO Programs have had little impact but a new survey of the field reveals a very different story”, available at: www.oliverwyman.com/media/OW_ENG_HLS_PUBL_The_ACO_Surprise.pdf (accessed April 18, 2015).


Further reading


About the authors

Peter F. Martelli is Assistant Professor of Healthcare Administration in the Sawyer Business School, Suffolk University, and is a Board member of the UC-Berkeley Center for Catastrophic Risk Management. His research focuses on intersections of theory and practice, with a focus on evidence-based management, organizational change, and behavioral models of risk and resilience, such as high reliability organizing (HRO). Martelli received the PhD degree in Health Services and Policy Analysis from the UC-Berkeley and an MSPH degree from Thomas Jefferson University, and was a Postdoctoral Fellow at the VA Center for Organization, Leadership and Management Research. He previously worked in the Scientific Policy Department at the American College of Physicians. Peter F. Martelli is the corresponding author and can be contacted at: pmartelli@suffolk.edu
Peter E. Rivard is Associate Professor of Healthcare Administration in the Sawyer Business School, Suffolk University in Boston. Prior to obtaining the PhD degree in Management from Boston College, he worked in healthcare management for 15 years and received an MHSA degree from the University of Michigan School of Public Health. His research explores quality, patient safety and performance improvement in ambulatory and primary care settings, primarily from the organizational learning, high reliability and leadership perspectives. He currently teaches leadership, ethics, research methods, and ambulatory and primary care management.

Karlene H. Roberts is Professor Emerita at the Haas School of Business, University of California, Berkeley. She received a Bachelor’s degree in Psychology from Stanford University, and the PhD degree in Psychology from UC-Berkeley. Roberts is a Fellow of the American Psychological Association, the American Psychological Society and the Academy of Management, and was conferred an honorary doctorate from Université Paul Cézanne, Aix-Marseille III. In 2011, she received the Academy of Management’s Practice Impact Award. She currently serves on the National Academies of Science’s Marine Board. Her research is concerned with the design and management of organizations and systems of organizations, in which errors can have catastrophic consequences. Roberts has examined both organizations that have experienced catastrophe (e.g. BP) and those that engage in seemingly precarious endeavors but have managed to avoid error over long periods of time (e.g. US Navy aircraft carrier operations). Her research has been supported by National Science Foundation, the Office of Naval Research, the US Department of Transportation and the California Energy Commission.