Creating Safety in Primary Care Practice with Electronic Medical Records Requires the Consideration of System Dynamics

Timothy R McEwen1*, Nancy C Elder2, & John M Flach1
1335 Fawcett Hall, Department of Psychology, Wright State University, Dayton, Ohio, 45324, USA
2PO Box 670582 Department of Family and Community Medicine, University of Cincinnati, Cincinnati, Ohio, 45267, USA

ABSTRACT
Improvement in quality and safety in health care often depends on eliminating errors. Using examples from our research on the medical testing processes in primary care medical practices, we argue that designing safer systems requires us to move beyond frameworks that focus exclusively on error elimination to consider the broader system dynamics that include information loops that can be critical to the overall stability of the system. We focus on describing the nature of information coupling in relation to the constructs of essential friction, autonomation, and ecological interface design and how these can lead to more resilient systems. With the recent push in the United States to move towards electronic medical records (EMR), we conclude with suggestions for improving EMR systems based on these ideas.

Keywords: Healthcare Quality, Safety, Essential Friction, Autonomation, Ecological Interface Design, Electronic Medical Records,

INTRODUCTION: ON WILDFIRES AND ERRORS

The major difference between a thing that might go wrong and a thing that cannot possibly go wrong is that when a thing that cannot possibly go wrong goes wrong it usually turns out to be impossible to get at or repair.

–Douglas Adams, Mostly Harmless

Based on a long history of studying human performance in high risk environments (e.g., aviation and nuclear power), researchers have discovered that eliminating error and creating safety are not necessarily synonymous, and in fact can sometimes result in systems becoming less safe or have created new unexpected problems [1, 2, 3]. This discovery has also been found to be the case in the healthcare domain [4]. Additionally, previous research has noted that primary health care providers express the belief that an electronic medical record (EMR) will “fix” the problems and errors associated with the testing process [5], but have EMRs have still been found to be error prone in spite of the technological 'advantage' [6]. To understand this, it is useful to make an analogy to wildfire prevention and suppression and its relation to safe and healthy forests.

Careful use of fire has been employed for centuries to encourage biodiversity. Vegetation that is burned periodically maintains high species diversity, and frequent burning of surface fuels limits fuel accumulation, thereby reducing the risk of larger fires. However, in 1937, Franklin D. Roosevelt initiated a nationwide fire prevention campaign, highlighting the role of human carelessness in forest fires and culminating with the adoption of Smokey
Bear as the official mascot of the US Forest Service. The emphasis of prevention and suppression allowed branches, leaves, twigs, and dead undergrowth (i.e., fuel load) to accumulate in forests around the country. The increased fuel loads led to situations where wildfires were able to reach beyond the underbrush to trees and their canopies, dramatically increasing the damage done, both in terms of area burned and severity of the damage. This problem has been referred to cynically as the legacy of Smokey Bear [5]. Human interference in the natural forest system circumvented nature’s natural control loops that allowed the forest system to self-regulate [5]. Thus, in an effort to decrease wildfire destruction, the policy of eliminating all wildfires created the potential for greater destruction by failing to respect the forest life cycle dynamics. Hence, many forestry experts now endorse changing the question from ‘how to eliminate all wildfires’ to ‘how to improve forest health through effective management.’

By applying the concepts of organizational resilience [6], we propose that errors and safety found in electronic medical record (EMR) systems are in many ways analogous to the example of the forest life cycle and wildfires. Just as a policy that sought to eradicate all fires led to a brittle system with potential for catastrophe, an EMR system that claims to eradicate all variations from prescribed procedures (i.e., errors) will also lead to a brittle system. In order to appreciate this analogy, it is important to note first that we are using ‘error’ to refer to deviations from a prescribed procedure or rule. Second, it must be noted that neither fires nor deviations (i.e., errors) are ‘good’ or ‘bad’ in any absolute sense. Fires that eliminate fuel on the forest floor and/or that promote germination of new growth can be good in the long run, but bad to those who might have lost life or property in the blaze. Similarly, a deviation from prescribed rules can be good (e.g., an innovation or adaptation) or a deviation can be bad, when it leads to negative consequences (e.g., root cause error). Unfortunately, the distinction between positive and negative deviations can typically only be made in hindsight [7]. Thus, the point is not to design a system to promote error, but rather to appreciate that in complex work it can be dangerous to design with the goal to eliminate all deviations or all variability with respect to any fixed normative model. There is danger in assuming that there is a single ‘best’ way.

Just as forests depend (in part) on wildfires to remain healthy, most systems depend on intelligent adaptations (e.g., heuristics and shortcuts) by the people within them to function efficiently; people experiment and bend the rules to accommodate situations not anticipated in the original formulation of processes. This need to adapt to complexity is one of the reasons why strictly ‘rule-based’ automated systems have had such limited success in healthcare or other safety critical systems (e.g., aviation or process control) [1, 2, 3]. The ‘reliability’ of strict rule-based systems becomes brittleness in the face of a complex, changing environment, where many gaps in processing exist [8, 9]. In these contexts, a healthy system is one that can bend (i.e., take short cuts; use heuristics) and stretch the rules (i.e., explore hypotheses) in order to adapt to changing circumstances to become resilient to those unexpected or unforeseen events [6].

In order to accommodate positive variability (e.g., heuristics and experiments) the design of systems must consider the information coupling that will allow the system to learn and adapt. For example, the design should consider the need for information about the boundary conditions underlying heuristics (e.g., limiting assumptions) and about the consequences of experiments. This information coupling can function as feedback (information from the result of the process that informs earlier part of the process) to help people aware of the consequences of actions) or feed forward (process dedicated to preventing certain anticipated events) to help people anticipate future events. Thus, the goal of eliminating errors must be balanced by the goal to enrich the information coupling. Historically, initiatives to automate have often fixated on eliminating deviations from a single best way and in the process have ignored the information coupling resulting in ‘brittle systems’ full of ‘automation surprises’ [3]. With the recent push towards the adoption of EMR systems, one motivation for writing this paper is out of concern that this same trend toward a ‘one best way design philosophy’ is already evident in healthcare design [10, 11].
Preventing manmade forest fires was an admirable and important goal, but when it led to forestry management that included the suppression of all fires, then forests became more brittle, meaning that the inevitable fires would become deadly and damaging conflagrations. For this reason, we propose that the design of sociotechnical systems, including EMRs, ought to be concerned with building resilient systems rather than systems based upon the rule based, error exclusion philosophies that can lead to system brittleness [6]. We believe that the question must also change from ‘how to eliminate all errors’ to ‘how to increase safety through effective management’ – where increasing safety involves BOTH reducing errors AND enriching the information coupling so that the system can adapt to the inevitable surprises that will be encountered in complex work environments.

The purpose of this paper is to explore the dynamics of information coupling and system dynamics as found in primary care medical practices with the ultimate goal of bringing these issues into the public discourse with respect to the design of resilient healthcare and EMR systems. To accomplish this, we will provide a number of examples from our research on the testing processes of primary care medical practices that illustrate the importance of rich information coupling and will conclude with some suggestions for improving the design of such systems based on the lessons we had learned.

**MEDICAL TESTING DOMAIN OVERVIEW**

One important component of medical care is the management of medical tests. Laboratory tests are commonly performed in all medical fields (e.g., primary care physicians order laboratory tests on 29 – 38% of their patients) [12]. As a common denominator among the various fields in medicine, the testing process can serve as a common health care task to study the healthcare system.

A normative model of the laboratory testing process has been proposed as a series of steps that fall into three stages (see Figure 1) [12]. First, the pre-analytic stage consists of the steps (e.g., ordering, collection, shipping, record keeping) that occur up until the patient’s specimen is analyzed. Second is the analytic stage consisting of the actual analysis of the specimen. During the post-analytic stage, results are sent back to the medical office for review, diagnosis, and follow-up.

In order to compare this normative model to the realities of actual practice, we triangulated our research using a combination of naturalistic observations, patient record audits, and critical incident interviews at four primary care offices to better understand the testing process. More detailed accounts of the methods and findings from this research can be found elsewhere [13, 14, 15, 16], but we will describe the interview method here. This study received approval from the University of Cincinnati institutional review board (IRB).

![Figure 1. Normative Testing Model (adapted from Hickner et al, 2005)](image)

**Interview method**

An important step in understanding the dynamics of a system is to explore not just why errors occur, but also the factors that make the system work (e.g., how it is able to function well much of the time) [17]. To this end, we interviewed 17 employees from four family medical practices in the greater Southwestern area of Ohio. We purposefully sampled a range of roles at each site including medical assistants (MA), physicians, office managers, nurses, and clerical support staff. We used a semi-structured interview format with interview probes based on the Cognitive Task Analysis (CTA) and Applied Cognitive Task
Analysis (ACTA) frameworks [18, 19]. Interviewees were asked to share examples when they caught an error in the system as well as when an error occurred.

All the interviews were analyzed by close reading of the notes and transcripts by three researchers; the variety of their expertise (family medicine and human factors psychology) added to the trustworthiness of the findings. Each interview was read, coded for major findings, discussed and then re-read and discussed again. Themes related to error stories including how and when errors were recognized were then placed into information coupling models.

**Interview results**

From these interviews we identified 63 incidents that either reflected breakdowns in the normative process (26) or acts of safety (37) in which people intervened to avert or prevent a breakdown of the process. In both cases, the incidents reflected variation with respect to the normative process. The distinction, however, was the overall impact on the process with the breakdowns increasing the risks to patients and the acts of safety decreasing the risks to patients.

In the following sections, we will use selected cases to illustrate three constructs that illustrate the challenge to designing resilient systems that should be considered in the design of EMR systems: Essential Friction, Autonomation, and Ecological Interface Design (EID).

**Essential Friction**

Essential friction was introduced by Rochlin as an analogy between the stability of automated processes and the stability of moving physical bodies [20]. He aptly notes that without friction, "[w]e would live in an impossibly kinetic world in which the consequences of every action would persist and multiply to the point of insanity" [20, p 132]. Likewise, processes without some amount of friction have the potential for errors to cascade throughout the system without opportunity for correction. In practice, essential friction typically takes the form of checks and coordination within a process. For example, in the testing process, each transition from one person to another (e.g., from the physician to the MA) might be seen as friction relative to an EMR system that would allow the physician to submit the testing request to a computer that was directly linked to the laboratory and for the results to be reported directly back to the physician via the same electronic system. On one hand, the process could be quicker and there would be fewer opportunities for the request to be misplaced. On the other hand, there were several incidents where people in the process, discovered and corrected errors made by the physician (e.g., entering incorrect information or missing critical results).

In one example from our research, it was reported that one physician occasionally signed off abnormal test results as normal. However, because his MA had taken the time to review the physician’s work, she noticed these errors and would return them to the physician for correction. Additionally, she began to highlight the abnormal values before the results were sent to the physician.

Friction, like wildfires and process deviations, has a value that is context dependent. From one perspective, each transition from one person to another can be seen as inefficiency (i.e., friction or waste), but from the other perspective having multiple people engaged in the process provides a degree of redundancy that can prevent an error from propagating unchecked through the system (i.e., friction as a source of stability). The lesson for EMR systems is that designers should not focus exclusively on using the automation to minimize human engagement in order to improve efficiency of flow through the normative process stages (eliminating human error or eliminating friction). Designers should also consider possibilities to incorporate ‘checks’ (i.e., essential friction) into the process. These checks may include automated checks to detect more common errors (e.g., requiring double entry of critical data as a verification check) or the system may be designed to alert humans to stages in the process where an error might create risk for the patient (e.g., highlighting and requiring explicit acknowledgement from the physician to critical test results that reflect the
need for action on the part of the medical staff). Thus, it is important to consider both the positive and negative effects associated with human ‘friction’ and to consider how to retain the positive aspects of this friction with the EMR systems.

**Autonomation**

When it comes to designing EMR systems, we should strive for *autonomated* systems, not *automated* systems. The difference between the two can be illustrated by Toyoda Sakichi’s (father of Toyota Motors) auto-activated weaving machine [21]. Prior to Sakichi’s design, automated weaving machines continued to operate even after a thread broke. The result was that defective products were “automatically” produced until the defect was discovered by an operator and the automation was turned off to repair the fault. On the other hand, Sakichi’s autonomated weaver included sensors that could detect thread breaks and would automatically shut down the process so that no time or materials were wasted on defective products. Thus, an autonomated system monitors the quality of the process and adapts or interrupts the process when errors that impact quality are detected. This leads to the counterintuitive advice that the first step to reliable systems is to make it easy to ‘turn the system off’ when a problem is detected. However, interrupting the systems helps ensure that problems are detected and resolved before they cascade to create greater difficulties. In some sense, the sensors are a kind of friction that ensures that the system catches problems, so that the system does not mindlessly (and efficiently) produce poor quality products.

As an example of this, it was reported than an office nurse went on leave for a week and upon returning she found a stack of test requisitions in her box waiting to be faxed to the hospital for referrals. Some orders had waited the entire week. Nobody had faxed them for her in her absence because they were not sure what to do with them. In this case, the ‘system’ continued ‘automatically’ as the requisitions piled up in the office nurse’s box. Would EMR systems make an event like this, (e.g., test requests or results piling up in a queue unattended) more or less likely? In the case observed, the pile in the nurse’s box was an obvious sign that the normal process had broken down, yet nobody noticed or intervened. In an EMR system the growing queue of unattended test results might be hidden in the computer – invisible to anyone but the specific client (the nurse on vacation). On the other hand, in an autonomated EMR system, checks can be designed to detect whenever tests ‘pile-up,’ unattended, and these tests can be automatically rerouted or others can be alerted to the bottleneck.

In contrast to the example of the nurse on vacation, one university-affiliated office employed many part-time physicians, meaning that these physicians only work in the office once a week. An MA made it a point to collect the awaiting test results for these physicians and preview them, ensuring that critical or abnormal results would quickly be addressed. The behavior of the MA is an example of autonomation. Could such foresight – to accommodate potential conflicts arising as a result of schedules that are not synchronous with the normative testing process be designed into EMR systems? This might include multiple alerts to several people for test results that indicate a need for intervention – to ensure that the results are noted or automatic rerouting of information that is not attended within a set time period. Thus, autonomation is a form of adaptive control, where the quality of the process is monitored and deviations that have potential consequences to the overall quality of the process result in adjustments to the process (e.g., rerouting information).

**Ecological Interface Design**

The concept of Ecological Interface Design (EID) builds on earlier constructs such as Direct Manipulation and Direct Perception to emphasize the need to make complex processes ‘visible’ to human decision makers [22]. This construct emerged as a result of the problem of data overload and the opportunity of graphical interfaces. The problem is most easily seen in the context of Nuclear Power Plants, where the classical single sensor, single
indicator display design would typically overwhelm operators with alarms during a critical event. With the development of graphical displays, alternative representations have been develop that make the internal dynamics of nuclear processes visible to operators (e.g., the flow of mass and energy within the processes could be illustrated in relation to the physical constraints like mass energy balances that shape the flow). Thus, the principle of EID is to use the power of automation to design visualizations or representations that help people to manage complex information.

At base, most of the incidents that were reported could be attributed to the difficult in tracking the testing process. For example, one MA called the reference laboratory to ask about the location of a missing result and was told that there was no record of that patient’s specimen. The logbook used to keep track of the specimens that had been picked up by the courier showed that the specimen had left the office, but the specimen had never been entered into the laboratory computer. The patient had to return to the office to provide a second specimen. In another case, the MA called the laboratory about a missing result (needed for patient care) to find that the laboratory had indeed received the specimen but had failed to process it. The laboratory admitted that they had not noticed the error until the MA had called. In yet another case, the MA contacted the hospital laboratory and asked for a result to be faxed to the office. The lab said that they would fax the result over. The MA went back to other tasks. The results never got back to the MA. She was not sure if the result was ever sent or if it got lost in the office. She had to call again for another copy.

Thus, the challenge is whether EMR systems will make it more or less difficult to visualize the testing process. The automation can be a ‘silent’ partner, leaving the humans completely out-of-the-loop for most of the testing process. Or, the automation can be a ‘team’ player that enhances the coupling of the people with the process. For example, package delivery services like UPS are able to precisely track packages – and customers can monitor the progress of their package on the internet. Scanning the packages at each point in the logistical chain can be viewed as waste because of the extra work that it entails or it could also be recognized as a helpful mechanism that allows problems to be addressed quickly by clearly revealing the system status for many sets of eyes (UPS agents and customers). EMR systems provide similar opportunities, but we often do not see this reflected in design or the literature. EID suggests that the goal in the design of EMR systems should be to enhance the visibility of the testing process, so that the medical staff can quickly ‘see’ where in the process a specific test is at any time. One potential solution with EMR systems could be to have an input field for ordered tests to enter a time limit that would alert staff that too much time had elapsed between steps. If a test result needed to be addressed within a certain timeframe and had not been addressed, the system could kick that result to another physician while also sending notification to the ordering physician that the result had been forwarded to close the loop. This type of information tracking can do more than just help in the tracking of individual tests. It can also allow more global analyses of general patterns in the testing activities within a practice. For example, it should be possible to query the system to see the distribution of tests at all stages of the process at a particular time, or to compute the average time for tests to pass through a particular stage over a particular period. For example, this might be useful to determine a choke point in the process such as an inefficiency in a lab or an unreliable courier service.

**CONCLUSION**

Given that system errors still occur despite having been “fixed” with EMR, it becomes more essential that we now incorporate the lessons learned from healthcare, aviation, and other industries into the design of EMR systems [24, 25, 1, 2, 3, 4]. Automation, designed around normative open-loop images of a process, can appear to improve efficiency and to eliminate the possibilities of human error. However, these systems often increase the ‘distance’ between the process and key human decision makers. In essence, the processes become hidden or buried inside the automation, preventing detection and correction of
errors and allowing small errors to cascade in ways that threaten system health. An explicit goal in the design of any EMR system should be to make the testing process, and other healthcare processes, as visible as possible – to enrich the information coupling. The goal should be to make it easier for the critical decision makers to ‘see’ where a test is in the process and to verify that the process is working as intended. This is the key to improving safety. With EMR systems it should be possible to ‘track’ and ‘see’ aspects of the testing process that had been invisible in paper systems. EMR systems make it possible for the office staff to ‘track’ the progression of the test after it leaves the office and to ‘see’ the progress relative to normative expectations (e.g., critical time windows).

In sum, we think it is a mistake to think of EMR systems as automated data management systems that are intended to displace the work of humans, to eliminate error, and to increase efficiency (in other words, to eliminate friction). Rather, EMR systems should be thought of as data visualization systems that are intended to enhance human perspicacity. The goal is not to stamp out all the fires, but rather to enhance the capacity to manage the health of the forest by improving the information coupling. Reducing human workload, eliminating error, and increasing efficiency are desirable ends for the design process. But these ends must be pursued in the context of the complex work dynamics. Ultimately, these ends are more likely to be achieved when the focus of the design effort is on enhancing the information coupling in ways that help people to manage the complexity.

ACKNOWLEDGEMENTS

This research is supported by a grant from the Agency for Healthcare Research and Quality: 1 K08 HS013914-01A2 to Nancy Elder, MD.

REFERENCES


