Challenge Project Part 3: Final

Electronic Health Record Usability Evaluation and Rating

David M. Schlossman, M.D.

Northwestern University

Medical Informatics 407

Winter 2012

Objectives

Based on preliminary data that proper implementation of health information technology (health IT) can reduce medical errors, improve patient safety, reduce waste and duplication, support evidence-based medicine, and reduce the overall cost of healthcare, the HITECH provisions of the American Recovery and Reinvestment Act of 2009 established large financial incentives for healthcare providers to adopt and demonstrate meaningful use (MU) of electronic health records (EHR). However, the majority of studies demonstrating successful effects of EHR's originated at just four institutions using internally developed EHR's, crafted and polished over 25 years with extensive input from their respective clinical staffs: Brigham and Women's Hospital in Boston, LDS Hospital in Salt Lake City, Vanderbilt Medical Center in Nashville, and Regenstrief Institute in Indianapolis. Attempts to reproduce those results using current generation commercially available EHR's have shown far less beneficial effect (Classen and Bates, 2011). In fact, the rush to implement EHR's and receive meaningful use incentive funds has overwhelmed EHR vendors' normal innovation and development cycles, forcing clinicians to use poorly designed information technology with significant negative effects such as increasing clinician fatigue and error rates as well as decreasing clinical productivity. A recent landmark study by the Institute of Medicine (IOM) Committee on Patient Safety and Health Information Technology concluded that, "designed and applied inappropriately, health IT can add an additional layer of complexity to the already complex delivery of healthcare which can lead to unintended adverse consequences" and "poor user interface design, poor workflow, and complex data interfaces are threats to patient safety (Warden et al., 2012)."

In addition to the problem of unintended adverse consequences, the MU incentives have not been optimally successful in driving physician adoption of EHR's. Preliminary data from the 2011 National Ambulatory Medical Care Survey shows that about 57% of office-based

physicians have made some use of electronic medical record/electronic health record technology and about 34% of physicians have an EHR system that meets the full criteria for a basic system (Hsaio et al., 2011). A 2009 study by Dr. David Blumenthal and colleagues from the Harvard Institute for Health Policy (Jha et al., 2009) showed that as of 2009 only 17% of U.S. hospitals surveyed had implemented computerized provider order entry for medications, only 7.6% had a basic electronic medical record in place, and only 1.5% had a functioning comprehensive EMR.

Studies of the barriers to EHR adoption (Boonstra et al., 2010; Castillo et al., 2010) have frequently identified poor usability as one of the most important barriers to EHR acceptance and use by physicians. The report of the Health Information Management Systems Society (HIMSS) EHR Usability Task Force (Belden et al., 2009) summarized the situation nicely:

> *We submit that usability is one of the major factors-possibly the most important factor-hindering widespread adoption of EMR's. Usability has a strong, often direct relationship with clinical productivity, error rates, user fatigue and user satisfaction-critical factors for EMR adoption. Clinicians lose productivity during the training days and for months afterwards as they adapt to the new tools and workflow.*

The usability problem represents a primary obstacle to both EHR adoption and to achieving many of the healthcare improvement goals that can only be accomplished through the use of health IT (ONC, 2011 p. 13). The Food and Drug Administration (FDA) has been tasked to collaborate with the National Institute of Standards and Technology (NIST) and the Agency for Healthcare Research and Quality (AHRQ) to "develop best practices to address systematic evaluation of usability with regard to patient safety" and to "provide perspectives on what constitutes usability and how to systematically improve the usability of EHR's (ONC, 2011).”

In 2011 NIST published draft guidance on the technical evaluation, testing, and validation of the usability of electronic health records (Schumacher et al., 2011). Like many other usability testing methods described in the literature (Belden et al., 2009; Rogers et al., 2009; Smelcer et al., 2009; Jaspers, 2009), the NIST procedures involve long complex testing

sessions which require a highly trained tester interacting with a clinician subject throughout the procedure. Such systems place a significant scheduling and work burden on the test subject and will likely deter participation by the busy physicians whose input is most needed. As the HIMSS comments on the NIST draft guidance point out “HIMSS believes that a critical missing factor in the draft guidelines is the inclusion of individuals with practical clinical experience… HIMSS suggests that testing should not be done in isolation and should represent typical care scenarios (Underwood and Lieber, 2011)." Common sense would suggest that the people who can best define and test usability are the primary EHR users, full-time clinical physicians and other healthcare providers. Yet this stakeholder group is severely underrepresented in EHR usability research and testing (Steele, 2009; Underwood and Lieber, 2011).

The fifth key area in the Startup America Policy Challenge (2012) is "unleashing market opportunities in industries like healthcare IT" and asking "the American public how to knock down barriers to innovation in healthcare IT." Regarding the usability barrier, the author proposes a low cost, customizable application for evaluating, rating, and comparing the usability of office and hospital EHR products. The application will use a testing presentation sufficiently straightforward and standardized to be presented as a website or on a laptop or tablet computer which can be used without the presence of a testing professional, at a time convenient for the physician, in an environment containing “relevant adverse factors such as interruptions, medical emergencies, competing workflow and time pressures that are more typical in a practice scenario (Underwood and Lieber, 2011).” Specific project objectives include:

- Identify and correct EMR design weaknesses that can lead to user errors and threats to patient safety.
- Motivate large numbers full-time clinical physicians to participate in ongoing testing and review of EHR usability.
- Accumulate reproducible quantitative data to guide vendors’ understanding of how to improve product design in this area.

- Inform the purchasing decisions of clinicians and healthcare organizations so that competition will motivate EHR vendors to improve usability performance.
- Stimulate the development of more logical, powerful, intuitive EHR's which support clinician workflows, genuinely increase clinician efficiency and productivity, and thereby improve the safety and quality of patient care.
- Overcome the usability barrier to EHR adoption

Scope Statement

According to the International Organization for Standardization (ISO), "Usability is the effectiveness, efficiency and satisfaction with which specific users can achieve a specific set of tasks in a particular environment (Schoeffel, 2003)." The HIMSS EHR Usability Task Force (Belden et al., 2009) developed this principle into five metrics which can be used to objectively evaluate usability and measure user performance in an EHR application:

- Efficiency, how fast a user can accomplish a given task.
- Effectiveness, how accurately a user can accomplish defined task goals.
- Ease of learning, the time and effort taken to reach a specified proficiency level.
- Cognitive load, the ability to organize data by meaningful relationships and optimize it for clinician thought processes.
- User satisfaction, the perceived effort necessary to complete important tasks.

To measure these parameters, the proposed application will leverage recently developed automated mobile usability analysis tools capable of recording, organizing, and reporting every detail of the interaction between a user and an application including all mouse clicks, timings, desktop activity, screen appearances, and a full audio and video record of the user at work (e.g. MORAE, 2012 and see also Bastien, 2010 p. e20). Such systems are also capable of administering user satisfaction questionnaires at the completion of the testing session. Important events can be flagged as they occur and metrics can be calculated automatically, dramatically decreasing the work required for analysis. Research so far has shown that usability testing with

automated data collection systems is an effective alternative to the more traditional laboratory-based testing methods (Bastien, 2010). As physician users work a set of commonly encountered clinical scenarios, the analytic tools will be used to measure the time necessary to complete important tasks (efficiency), the accuracy in completing important tasks, (effectiveness), and the perceived effort necessary to complete important tasks (satisfaction). Although well validated methodologies and guidelines exist to measure ease of learning (Grossman et al., 2009) and cognitive load (Embrey et al., 2006), these measurements are more complex and will be deferred to the second generation of the application.

For use in the scenarios, the application designers will construct a database of 100 imaginary test patients with data available on each patient to include demographics, payer information, problem list, allergy list, medication list, and flow sheets of representative laboratory data. The database will be constructed in Oracle, Microsoft SQL, and comma delimited (CSV) text formats, which should allow sufficient flexibility for import into most testing systems. A similar database was constructed in the author's own practice for training use when the current office EHR was implemented in 2007, so this is known to be feasible. The design team will also construct a library of clinical scenarios representing important tasks clinicians commonly wish to accomplish using EHR's. The HIMSS and the National Institute of Standards and Technology (NIST) have published model scenarios and enumerated the EHR-specific tasks involved in completing each of them. A typical outpatient scenario and task list from the HIMSS (Grayson et al., 2010, p. 17) would be:

New patient presents with 3 days of dysuria, hematuria, urgency and frequency. No fevers, chills, or back pain. Physician orders Bactrim DS 1 tab bid x 3 days, phenazopyridine 200 mg tid after meals. Two days later, patient phones in that she has an itchy rash between the toes of her left foot. Physician recommends OTC terbinafine cream to apply bid x 10 days.

A typical inpatient scenario and task list from the NIST (Schumacher et al., 2011, p. 82):

Patient is a 45-year-old African-American female living in an urban center. She has hypertension (HTN), obesity, Type 2 Diabetes, elevated cholesterol (LDL), and asthma. She started smoking when she was 17 years old and is actively trying to quit. The patient is brought to the Emergency Room by Emergency Medical Services. She called 911 when she was having chest pain.

Using similar techniques, the design team will initially construct a panel of 10 inpatient and 10 outpatient scenarios with associated structured lists of the tasks necessary to complete each one. As clinicians log in to the EMR under study and work the scenarios, the monitoring application will capture the time for each task, time for the whole scenario, and task accuracy as indicated by number of correct clicks, use of the back button, number of pauses to decide what to do next, and the number of tasks completed successfully. At the end of a session, the user will complete a questionnaire, ranking several aspects of the EMR system's ability to support clinical task completion and decision making on a scale of one to five and allowing the calculation of a satisfaction score. An example of such a survey form can be found in Grayson et al., 2010, p. 25-26. The application will keep a complete audio and video record of each session, and users will be encouraged to "think aloud" and comment during testing. The application will also provide statistical tools to aggregate and analyze the data from multiple sessions and customization tools for end-users to add test patient data, clinical scenarios, and survey instruments to make the program more useful and effective at their respective institutions.

Resource Needs

The proposed application could easily reside and run on an Intel core i5 based laptop computer with 8 GB of RAM, a one TB hard drive, a webcam, a microphone, and Windows 7 operating system, a configuration that can be obtained for less than $800. The MORAE recording and analytics software bundle costs $1,495, and additional copies of the recorder module are available for $195. Discounts are available for government, nonprofit, and educational organizations. This laptop configuration could be utilized for testing anywhere there

is a connection to a hospital or physicians' office computer network or even over the internet (Jaspers, 2009). Additional work to complete the application will include constructing the sample patient and scenario libraries and task lists, writing the user satisfaction survey documents, writing the software and interface to present user instructions, help, and survey instruments, and writing the software to package and store test session results in a standard format which allows the results of multiple testing sessions to be combined and statistically analyzed. A team of three experienced physicians working half-time and three full-time software development engineers should be able to produce a working prototype application in six months. Time will be required for qualification and recruitment of test participants, for alpha and beta testing and for validation testing in comparison to more established usability evaluation systems. Even with the additional development and testing time, a viable release candidate could be produced within a year. An estimated project budget for year one would be as follows:

Line Item	Cost
Computer Hardware	$3,500
Computer Software	$4,000
Physician Salaries (1.5 FTE)	$300,000
Software Development Engineer Salaries (3 FTE)	$270,000
Recruitment and Testing	$22,500
Total	$600,000
Median Physician and Software Engineer salaries retrieved from United States Bureau of Labor Statistics www.bls.gov/oes/current/	

Risk Management

Because the application is intended as an end-user product which tests but does not alter EMR systems, it should not necessarily be subject to any laws or government regulations. The

construction and use of a model patient database is specifically designed to keep the entire application in a testing environment and avoid any possible application of HIPAA regulations. The NIST draft guidance on usability evaluation published on September 28, 2011 (Schumacher et al., 2011) and the proposed Stage 2 MU Rule published and opened for comment in the Federal Register on March 7, 2012 (DHHS, 2012) both address federal government mechanisms for EHR usability testing. With the caveat that time for review has been extremely short, the author's initial impression is that the MU Regulation's approach to usability is still somewhat nonspecific, and as noted above, the NIST procedures are very cumbersome. Both of these processes are directed at EHR analysis and certification during product development. The proposed application will largely be applied after EHR's are released in the market and will be more easily customizable to evaluate specific local EHR implementations in their operating environments. This additional functionality has the potential to identify additional usability issues missed during the certification process and will therefore complement rather than oppose federal government certification systems. As Classen and Bates (2011) point out:

> *Our health care system needs tools for evaluation of these systems when they are operational, not just before implementation…Such self-assessment tools could be developed and implemented with the use of simulation approaches similar to the way the Leapfrog Group assessment tool for EHR's and clinical decision support has been used.*

Ethical Guidance

The author believes that no significant ethical dilemmas will arise from the deployment of the proposed application. However, ethical issues played a significant role in the origin of the project. The HIMSS EHR Usability Pain Points Survey (Ribitsky et al., 2010) found medium and high severity usability problems to be very frequent among physician and nurse users of EHR's. The AHRQ Electronic Health Record Usability: vendor practices and perspectives report (McDonnell et. al., 2010) found that "formal usability testing, the use of user-centered design processes, and specific resource personnel with expertise in usability engineering are not

common" and that "many vendors did not initially address potential negative impacts of their products as a priority design issue." More than 90% of the membership of many federal advisory committees such as the HIT Policy Committee is composed of industry stakeholders, consumer advocates, academic and administrative physicians, non-physician administrators and representatives of third-party payers, employers, labor groups, etc. (Steele, 2009). The voice of the busy full-time clinical practice physicians who deliver approximately 80% of the patient care in the U.S. has difficulty being heard on this issue. One goal of the proposed application is to give this group additional *autonomy*, to help them make their usability concerns objective, concrete, and actionable as a first step in solving the problem. Also, as the IOM report (Warden et al., 2012) demonstrates, failure to correct these problems can present an imminent threat to patient safety, so efforts to make progress towards a rapid solution are consistent with the *do no harm* principle of ethics.

Stakeholders' Roles

Although the author does not believe any government agency has regulatory authority specific to the proposed application, he favors continuing the use of regulatory authority to require measurement and certification of minimum usability standards in EHR products to provide a stronger incentive for EHR vendors to truly incorporate user-centered design principles in their products and develop expertise in this area. The large volume of informatics research and the amount of attention from professional societies devoted to this problem suggest it is of great concern to healthcare organizations, providers, and medical informatics professionals alike. The natural "consumers" for this application are physicians, nurses, Allied health professionals, and healthcare organizations. They are likely to embrace a product that can help develop make electronic health records more simple, natural, consistent, forgiving, and supportive of workflow (Belden et al., 2009). It would be expected that a product designed to reduce medical errors,

improve patient safety, and improve physician efficiency thereby decreasing costs should also be strongly favored by patients and patient advocacy groups. There likely will be significant opposition from EHR vendors. Data from the proposed application, new demands from their customer base, and new federal regulations will add complexity and expense to their software development processes. We all live in a world where our professional performance is judged against relevant objective external standards and benchmarks. For physicians this means accepting and embracing clinical pathways, evidence-based best practices, and external monitoring of the quality and cost-effectiveness of our clinical work. For EHR vendors it means recognizing that their products do not operate in a vacuum, but are part of a larger sociotechnical system where the quality and performance of the product is strongly linked to what healthcare providers can accomplish when they use it. If the vendors are truly committed to using their technology to improve health care safety, quality, and efficiency, they will be open to cooperation with clinicians and participation in an iterative development process of usability testing and improvement leading to overall better EHR's.

References

Bastien, J.M.C. (2010). Usability testing: a review of some methodological and technical aspects of the method. *Int J Med Inform 79*: e18-e23. Doi: 10.1016/ijmedinf .2004.12.004.

Belden, J.L., Grayson, R., and Barnes, J. (2009). Defining and testing EMR usability: principles and proposed methods of EMR usability evaluation and rating. Report of the HIMSS EHR usability task force. Retrieved from http://www.himss.org/content/files/himss_definingandtestingemrusability.pdf.

Boonstra, A. and Broekhuis, M. (2010). Barriers to the acceptance of electronic medical records by physicians: from systematic review to taxonomy and interventions. *BMC Health Services Research, 10*:231-248, doi:10.1186/1472-6963-10-231.

Castillo, V.H., Martinez-Garcia, A., and Pulido, JRG. (2010). A knowledge-based taxonomy of critical Factors for adopting electronic health record systems by physicians: a systematic literature review. *BMC Medical Informatics and Decision Making, 10*:60-77, doi:10.1186/1472-6947-10-60.

Classen, D.C. and Bates, D.W. (2011). Finding the meaning in meaningful use. *N Engl J Med 365*(9): 855-858.

Department Of Health And Human Services (DHHS) (2012). Medicare and Medicaid Programs: Electronic Health Record Incentive Program-Stage 2. *Federal Register 77*(45): 13698 Retrieved from http://www.gpo.gov/fdsys/pkg/FR-2012-03-07/pdf/2012-4443.pdf.

Embrey, D., Blackett, C., and Peachey, J. (2009). Development of a human cognitive workload assessment tool. Retrieved from http://www.mcga.gov.uk/c4mca/research_report_546.pdf.

Grayson, R., Aranello, P., Bochinski, J., Boone, E., Phua, J., and Uehara, R. (2010). Selecting an EHR for your practice: evaluating usability. Report of the HIMSS usability task force-EHR selection guide work group. Retrieved from http://www.himss.org/content/files/HIMSS%20Guide%20to%20Usability_Selecting%20an%20EMR.pdf.

Grossman, T., Fitzmaurice, G., and Attar, R. (2009). Survey of software learnability: metrics, methodologies, and guidelines. Proceedings of the 27th International Conference on Human Factors in Computing Systems. New York, NY: CHI '09.ACM pp. 649-658.

Jaspers, M.W.M. (2009). A comparison of usability methods for testing interactive health technologies: methodological aspects and empirical evidence. *Int J Med Inform, 78*(5): 340-353.

Jha, A.K., DesRoches, C.M., Campbell, E.G., Donelan, K., Rao, S.R., Ferris, T.G., Shields, A., Rosenbaum, S., and Blumenthal, D. (2009). Use of electronic health records in US hospitals. *N Eng J Med, 360*(16):1628-38.

Hsaio, C.J., Hing, E., and Cai, B. (2011). Electronic health record systems and intent to apply for meaningful use incentives among office-based physician practices: United States, 2001-2011. *NCHS data brief* No. 79, Hyattsville, MD: National Center for Health Statistics.

McDonnell, C., Werner, K, and Wendel, L. (2010). Electronic health record usability: vendor practices and perspectives. Rockville, MD: Agency for Healthcare Research and Quality Publication No. 09(10)-0091-3-EF.

MORAE (2012). http://www.techsmith.com/morae.

Office of the National Coordinator for Health Information Technology (ONC) (2011). Federal Health Information Technology Strategic Plan 2011-2015. Retrieved from http://web.mediacdt.com/onc-emerg/FINAL-Federal-Health-IT-Strategic-Plan-0911.pdf.

Ribitzky, R., Sterling, M., and Bradley, V. (2010). HIMSS EHR usability pain point survey Q4-2009. Retrieved from http://www.himss.org/content/files/Usability_Pain_PointsHIMSS10.pdf.

Rogers, M.L., Patterson, E.S., Chapman, R.J., and Render, M.L. (2005). Usability testing and the relation of clinical information systems to patient safety. In Henriksen, K., Battles, J.B., Marks, E., and Lewin, D.I. (Eds) *Advances in patient safety: from research to implementation.* Volume 2, Rockville, MD: Agency for Healthcare Research and Quality Publication Nos. 050021 (1-4) pp. 365-378.

Schoeffel, R. (2003). *ISO Bull 34*: 6-7.

Schumacher, R.M., Patterson, E.S., North, R., Zhang, J., Lowry, S.Z., Quinn, M.T., and Ramaiah, M. (2011). NIST draft guidance on technical evaluation, testing, and validation of the usability of electronic health records. Retrieved from http://www.nist.gov/healthcare/usability/upload/Draft_EUP_09_28_11.pdf.

Smelcer, J.B., Miller-Jacobs, H., and Kantrovich, L. (2009). Usability of electronic medical records. *J Usability Stud, 4*(2): 70-84.

Startup America Policy Challenge (2012). http://challenge.gov/White%20House/257-startup-america-policy-challenge.

Steele, E. (2009). The Voice of the Physician. EHR Straight Talk. Retrieved from http://blog.srssoft.com/2009/02/the-voice-of-the-physician/.

Underwood, C.S. and Lieber, H.S. (2011). Letter to Patrick D. Gallagher, Director, NIST commenting on NIST draft guidance on technical evaluation, testing, and validation of the usability of electronic health records. Retrieved from http://www.himss.org/content/files/EHR_Usability/20111110_HIMSS_Response_NIST.pdf.

Warden, G.L. (Chair) and the Institute of Medicine Committee on Patient Safety and Health Information Technology (2012). *Health IT and patient safety: Building safer systems for better care*. Washington, DC: National Academies Press.