INFORMATICS PROFESSIONALS. LEADING THE WAY.

November 10, 2011

Patrick D. Gallagher
Under Secretary for Standards and Technology and Director
National Institute of Standards of Technology
U.S. Department of Commerce
1401 Constitution Ave NW
Washington, DC 2023
Sent via email to: EHRUsability@nist.gov.

Re: Draft Guidance on Technical Evaluation, Testing and Validation of the Usability of Electronic Health Records (NISTIR 7804)

Dear Mr. Gallagher:

On behalf of the American Medical Informatics Association (AMIA), I am pleased to submit these comments in response to the above-referenced Draft Guidance. As health care professionals and informaticians, many of our members are involved in biomedical and health informatics research and evaluation as well as application development and implementation. AMIA is deeply interested in the Draft Guidance as it pertains to a long standing and growing area of informatics and health information technology. We welcome the opportunity to contribute to the public commenting process about a critical but complex topic.

We applaud NIST's efforts to highlight the importance of usability testing for electronic health records (EHRs) through its issuance of proposed guidance regarding usability, evaluation, and electronic health records. Overall AMIA believes that the draft Guidance provides very practical information to support usability evaluation of EHRs. The proposed three step process offers a basic framework for considering the general steps involved in producing useful, usable, and satisfying interfaces to EHRs. We believe that the descriptions of formative usability testing and summative validation, principles for selecting representative samples for testing, and definition of terms are potentially helpful to a diverse community of users, developers, and vendors. We are pleased that the Guidance offers a baseline for applying basic usability testing in accordance with minimum standards for formative and summative evaluation. However, AMIA would like to offer the following comments and suggestions:

Readability

AMIA believes that NIST should include a more comprehensive discussion of the intended scope of and audience for the Guidance. We suggest that revisions are needed to assure that the Guidance is simpler, more readable, and thus understandable. The Guidance contains potentially valuable material that has not been fully synthesized, so it appears disjointed, repetitious, and very hard to follow in some sections, particularly in Section 5 (see below). There are a few places where additional detail and explanation would be extremely helpful (e.g., reporting results.)

Proposed Users

We are concerned about the limited and narrow identification of EHR users. We believe that this is a significant oversight and has implications for broad user acceptance of the Guidance. The Guidance describes users as physicians, physician assistants, advance practice nurses, registered nurses and licensed practical nurses. Since user testing by role is a key component of thorough formative and summative testing in laboratory environments as well as in live-production environments, we urge NIST to include all ancillary professional roles in multidisciplinary teams across diverse settings (acute care and ambulatory care). These roles include: pharmacists, respiratory therapists, physical therapists, occupational therapists, dieticians, departmental technicians (e.g. EKG tech), and a variety of trainees, including medical and nursing students.

However, user testing alone (after systems development) is insufficient to ensure that systems reflect user needs (re: expert reviewer section page 27). As one AMIA member stated, “I was recently told by an EHR vendor that all they need to build good systems is to have undergraduate engineers shadow practitioners for a few weeks to learn their needs.” This practice has no doubt contributed to the building of EHRs that do not work well for clinicians. Clinicians have knowledge about their clinical practices, care delivery, and workflow that computer technicians do not possess and cannot glean through a few weeks of observations and limited usability tests. Failure to recognize and include the multidisciplinary health care (provider) team along with technical (biomedical engineers, computer scientists) and clinical domain personnel (biomedical and health informaticians) also has serious ramifications for patient safety and quality of care.

The statement on page 9 in italics states that "physicians" have trouble finding information - this is an overly narrow statement, since EHR's are used by the entire multidisciplinary healthcare team. Also the Guidance does not reflect the increasing role and engagement of the patient and family as part of the team. We are concerned that in general the personnel identified in the Guidance do not sufficiently understand the nuances of the healthcare domain. We urge NIST to

recognize that usability testing and evaluation must be performed by individuals who understand the tasks (both cognitive and physical), people, clinical and administrative work flows and processes, and "communication" techniques common in delivering patient care.

Proposed Evaluative Framework

We believe that the Guidance narrowly focuses and over-emphasizes support of current Federal efforts to implement Meaningful Use (MU). We urge NIST to reassess and revise the Guidance to assure that it can be applied within broader EHR implementation efforts while still supporting current MU efforts.

We believe that the proposed protocol will be useful to evaluate design and implementation errors that can evoke patient safety and medical errors. However, we are concerned that the proposed approach is a narrow subset of a much bigger field, science and knowledge base. This broader field includes attention to multiple factors such as ease of use, intuitiveness, match with a user's mental model of the domain, clarity of workflow and navigation, ease of data capture and output, and cognitive supports in display design. Many human factors and safety design features affect an end user's adoption, willingness, and ability to use a system. Providers continue openly to voice concerns about EHR implementation challenges (such as efficiency of work flow and degradation in work load) that present obstacles to delivering safe and quality care. We believe that consideration of more comprehensive usability factors is critical to moving the industry beyond current user perceptions that EHR systems are burdensome and unhelpful to providers.

Testing environments

We are concerned that the Guidance does not recognize the differences between live production environments of users ("naturalistic") from the laboratory instance of a given vendor's test environment. While the draft requires the laboratory to be set up as a distraction free environment so that "performance issues are not attributed to exogenous environmental factors", we believe that in the "real world" the conditions under which EHRs are deployed are filled with excessive audio and visual distractions, and multiple interruptions. To test without "realistic" conditions would not provide useful results to inform actual usability.

While AMIA recognizes that real world testing may pose logistical and practical challenges given the complex and dynamic processes of EHR system design and development, we believe that testing is required in both laboratory and real world environments. In particular, real world usability testing can be used to help product developers enhance and improve subsequent product

releases, thus assuring the ongoing usability of the product. Further, AMIA believes that changes are inevitable regarding EHR product design and development; thus, subsequent product releases must be also be subject to usability testing and evaluation.

Moreover, because of the already large outlay of funds in acquiring and implementing an EHR system, organizations will continue to support fixes rather than discard a poor system – though fixes to a poor system over time are likely to far exceed the cost of removing and replacing the existing system. Also, when modules or parts of EHR systems are updated, testing of a "component" in isolation is insufficient to demonstrate appropriate usability. What is needed is testing that validates the utility of the new feature or component within the larger EHR under real time use.

At the heart of all safety and quality performance is the culture of the institutions and people who work within the organization or work unit. Testing for accuracy of data entry and clarity of information display is not sufficient. The ultimate goals of the system as a tool to help provide care to patients must always be kept at the forefront. One AMIA member puts it this way, "Unless EHRs allow doctors, nurses, clerks, and other professionals the needed support to think about what they are doing, so that they can make corrections based on their own intellect, we are taking a step backwards."

We believe that it is appropriate to study the software process from design through development, but it is critical that the study not stop with pre-defined test scripts run against a system in a laboratory environment. Unless testing is done at a site where the EHR is actually implemented, the testing tasks will not be sufficiently realistic. No matter how thorough the design and development processes are, it is very difficult to identify all of the challenges that can arise when actual users attempt to use the system to deliver care in an actual practice setting. Users are likely to encounter situations not anticipated or predicted by developers. Furthermore, if users attempt to address challenges in the real world, some may choose to implement a work-around that can have devastating consequences. Also, implementation of a system amid all of the complexities of actual clinical practices often results in configuration challenges unanticipated by developers.

Ongoing Usability Testing and Evaluation throughout the System's Life Cycle

AMIA urges NIST to look beyond laboratory based usability testing prior to system go-live. There needs to be guidance for ongoing testing that occurs in the user environment after system

implementation, since some problems, particularly those which may be a threat to patient safety, will arise sporadically and only after routine use is underway.[1]

On page 8, the Guidance states, "It is our expectation that the potential for all of these use errors can be identified and mitigated based on a summative usability test conducted by qualified usability/human factors professionals prior to EHR implementation/deployment." Although tests prior to implementation are likely to identify a substantial proportion of usability issues, it is particularly important to identify issues related to clinical workflow. In analyses of incidents reported to the FDA, researchers found that mismatches with clinical workflow were a major contributor to adverse events. [2] [3] [4]

We are concerned that testing after development may encourage developers to focus only on that aspect of testing. Instead, formative usability testing across the systems life-cycle is needed. We urge NIST to include a section on user-centered design in the Guidance. Further, we believe that expert evaluation methods should include clinical usability as well as traditional (technical) usability experts. We believe that usability testing with actual users, including a complement of users from different disciplines, needs to be heavily weighted in addition to those undertaken with or by experts. The expert evaluations should occur earlier in the systems development life cycle, so that findings related to design can be incorporated into more robust prototypes.

Three-phase process

The Draft Guidance recommends a three-phase process for usability evaluation of EHRs. Step 1 recommends incorporation of users, work settings and common workflow into design. However, there is no clear recommendation for how this can be done nor is a proposed process included. While we agree that usability needs to be incorporated into the design process, it is not clear how

---

[1] Bloomrosen M, Starren J, Lorenzi NM, Ash JS, Patel VL, Shortliffe EH. Anticipating and addressing the unintended consequences of health IT and policy: a report from the AMIA 2009 Health Policy Meeting. J Am Med Inform Assoc. 2011 Jan-Feb;18(1):82-90.

[2] Magrabi F, Ong MS, Runciman W, Coiera E. Patient harm associated with healthcare information technology: an analysis of events reported to the US Food and Drug Administration. AMIA Annual Symposium; Washington DC2011

[3] Magrabi F, Ong MS, Runciman W, Coiera E. Using FDA reports to inform a classification for health information technology safety problems. J Am Med Inform Assoc. 2011.

[4] Kushniruk AW, Borycki EM, Kuwata S, Kannry J. Emerging approaches to usability evaluation of health information systems: towards in-situ analysis of complex healthcare systems and environments. Stud Health Technol Inform. 2011;169:915-9.

NIST is suggesting that this be accomplished. We believe that methods for examining healthcare workflows that go beyond the interaction of individual clinicians with an EHR are needed. The process of patient care needs to be examined with a multidisciplinary and team-based focus so that users can adequately examine how EHRs impact the overall process of care. This is arguably an area for ongoing study by researchers and academics which then needs to inform the design process in companies (which don't have the same access to human subject research environments). AMIA encourages stakeholders to leverage the significant contributions of usability and human factors research and science to enhanced EHR systems design and usability testing.

While the Guidance includes a recommendation for EHR vendors to have live, "near" production instances of their EHR system, including content and de-identified patient records, to perform full summative testing, we are concerned that this proposed method is not sufficiently robust to test how the product will be implemented in actual health organization's production environments.

The Guidance is really fairly silent about real-world testing. We encourage NIST to consider a fourth phase that includes real world testing of EHRs by diverse users at pre-determined intervals after implementation. We urge NIST to develop and recommend a more appropriate standard for how health care organizations can realistically test their systems before, during and after they are "deployed and implemented". Additionally, we strongly suggest that NIST further differentiate between specific guidance and protocols for performing formative and summative testing by EHR developers and vendors, from those usability testing methods and approaches for use by health care organizations. AMIA believes that user studies should not end within the simulation laboratory phase. Even with high fidelity simulations, it is impossible completely to replicate the dynamic and unpredictable nature of real world health care environments and associated worker responsive work flow in a lab setting. The proposed tightly controlled evaluations are likely to be inadequately representative.

Step III of the protocol states discussed summative evaluation. However, it is not clear if this is meant to be conducted as a formal research study. Further, the overall approach appears unrealistic and potentially too prescriptive. For example, inclusion criteria for participants are much too limited. Organizations need to be sure to include a wide range of potential system users (and teams of users) in ways that are appropriate to their roles. At a minimum this should include pharmacists, social workers, nutritionists, physical therapists and others in addition to physicians and nurses. AMIA believes that the organization and the development team need to decide who the appropriate mix of personnel is. In addition, we suggest that participants should be chosen based on domain and subject matter experience as well as their experience using an EHR. Those

who have expertise in particular clinical settings or in specialty settings (such as cardiac care or pediatrics) are likely to have different input than those who are novices to the setting.

While it is very important to conduct continued evaluations in all of the settings in which the EHR is used, the Guidance does not specify how or when this should take place. We believe that this could be emphasized as a clearly stated fourth step in the overall approach to testing.

In addition, AMIA believes that the use of qualitative methods needs further clarification. The Guidance should clearly articulate what information is important to focus on when conducting interviews and focus groups. Observations need to be structured in a way that ensures that the entire healthcare team including patients and families is taken into account in the analysis.

In the following discussion, AMIA offers additional comments by section and/or page number:

- Section 3.2, Step 1, Usability Analysis. The first bullet indicates the users should be described and indicates one category is "job classification." Several usability studies have indicated that role-based functions (e.g., clinical informaticians) are more helpful than job titles or professional categories in determining characteristics of users. For example, pediatricians and surgeons are known to have very different views of what is important to see early when accessing an EHR, so a generic "physician's view" would typically generate different reactions and use patterns from the two types of practitioners.

- Section 3.3 (pg 13) -what the usability protocol is not: It is not clear how the suggested protocol addresses the optimization of workflow and efficiency. It seems that the protocol is focused on addressing the identification of specific issues in the user interface that may lead to errors or threats to patient safety. This is an important part of usability evaluation but on its own it is not sufficiently comprehensive. If the Guidance retains this narrow focus, we suggest that the Guidance clearly indicate that it only addresses limited components of usability testing.

- Section 5.1, Figure 2 is not explained in the text and it's not clear what it contributes.

- On page 18 Figure 3 presents a model for analysis and understanding of use related risks of EHR systems. The basis for the proposed model is unclear, since no reference is provided. The classification for adverse events appears insufficient; NIST may wish to consider using an existing patient safety classification. For example the WHO

International Classification for Patient Safety uses the "degrees of harm" to examine actual and potential patient outcomes [5] [6]

- Section 5.1. Figure 3 is critically important (and far more useful than Figure 2), but its content is not parallel to the supporting text on pages 18 and 19. For example, the leftmost box has very specific types of problems while the bulleted list has categories that are more inclusive and more useful for a model.

- Section 5.1, page 21. The listing of adverse events does not seem clinically appropriate and the opening headings are not intuitive. NIST may wish to re-order the categories of adverse outcomes with the most frequent types of errors at the top, in descending order. With all the literature review in this Guidance, we assume that it should be possible to determine a better presentation of these events.

- Section 5.1 & 5.4. These sections describe medical errors and adverse events. If the Guidance is aimed at clinical implementers of EHR systems who need to understand usability then the details describing medical error appear to be unnecessary. However, if the Guidance is aimed at usability experts who do not understand the clinical issues related to the usability of an EHR, then perhaps this kind of detail is necessary.

- Section 5.2. This section organizes many of the same concepts as Section 5.1 in a different and better way, but because of the sequence it appears to be repetitive. The definitions of usability should appear earlier in the Guidance.

- Step II of the protocol - Expert review/analysis: pg 12 states that usability/human factors and clinical subject matter experts conduct the review. However, in section 6 which provides the details for conducting the expert review, there is no further mention or description of the role of the clinical subject matter expert in the review process. Clinical subject matter experts are clearly important to this process. Their roles should be clarified.

---

[5] Runciman W, Hibbert P, Thomson R, Van Der Schaaf T, Sherman H, Lewalle P. Towards an International Classification for Patient Safety: key concepts and terms. Int J Qual Health Care. 2009;21(1):18-26.

[6] Sherman H, Castro G, Fletcher M, Hatlie M, Hibbert P, Jakob R, Koss R, Lewalle P, Loeb J, Perneger T, Runciman W, Thomson R, Van Der Schaaf T, Virtanen M. Towards an International Classification for Patient Safety: the conceptual framework. Int J Qual Health Care. 2009;21(1):2-8.

- "Page 16 “Methods for causing change. Improving patient safety in an organization typically requires reducing gaps between acknowledged guidelines, standards or protocols and practice through multiple strategies, including standardization, monitoring relevant measures and collaboration across organizations.” It is not clear what NIST means by “standardization.”

- Methods for conducting the review include the assessment of evaluative indicators (page 20). It is unclear how these evaluative indicators are to be assessed. There is a recommendation for the use of qualitative methods to understand clinicians’ response to system use. It is not clear if this is meant to occur after the initial lab-based testing.

- Section 6.1. The opening list of bullets jumps back and forth between characteristics of users and systems. The second paragraph presenting the SHARPC grant information appears to be out of sequence.

- Section 6.1. Third paragraph, describing who the experts participating in the reviews should be. We wonder that the rationale is for naming different types of degrees in such detail? Is there any evidence that education is a better criterion than professional experience, which is listed twice and yet seems to be undervalued? (page 28, top)

- Section 6.2. The first paragraph in this section refers back to Figure 3, which is a conceptual model but does not offer the same framework as the Guidance in Appendix B. This appears to be an editing error or inconsistency in versioning.

- Section 6.2. The last paragraph indicates in one sentence that there should be more than one expert evaluator. The lack of guidance here is surprising and seems like an afterthought. The Guidance should be clearer on an optimal number, based on the review of evidence or expert opinion or a rule of thumb/ best practice, such as 3-5 experts.

- Section 7. Step 7 (page 47). The section “Analyze Data and Report the Results” is cursory and needs to be expanded. Appendices K and L do not seem particularly helpful in terms of guidance for reporting. Compared with the level of detail in other sections of the Guidance, the lack of attention to this section is surprising.

AMIA's Usability Task Force

AMIA has convened a multi-disciplinary Usability Task Force to further explore the complexities and challenges inherent in electronic health record usability practice – design, evaluation, and testing. The topic is not without controversy. Issues of regulation, certification, and meaningful use criteria have become entwined with the topic of EHR usability. The timely and critical importance of usability of EHRs is underscored by the number of projects underway by several public and private sector organizations in addition to those undertaken by NIST. We expect that the recommendations from the AMIA Usability Task Force will be extremely important for the field and for health care more broadly. In addition to addressing some of the questions raised in these comments, the AMIA Task Force will explore a number of topics including: the role that current and future research can and should play in usability science and the practical implications of incorporating usability testing into the product life cycle. We look forward to sharing the results of our Task Force's deliberations when it has completed its work in early 2012.

Summary

We appreciate the opportunity to provide these comments. AMIA thanks NIST for issuing this Draft Guidance and for its timely attention to an important technical as well as public policy issue. AMIA is a source of informed, unbiased opinions on policy issues relating to the national health information infrastructure, the uses and protection of clinical and personal health information, and a variety of public health considerations. We stand ready to assist NIST in addressing the topics noted above. Please feel free to contact us at any time for further discussion of the issues we have raised.

Respectfully submitted,

Edward H. Shortliffe, MD, PhD
President and CEO