Decision Support for Acute Problems: The Role of the Standardized Patient in Usability Testing

Jeffrey A. Linder\textsuperscript{a}, MD, MPH, Alan F. Rose\textsuperscript{b}, MSc, Matvey B. Palchuk\textsuperscript{b}, MD, MS, Frank Chang\textsuperscript{b}, MSE, Jeffrey L. Schnipper\textsuperscript{a}, MD, MPH, Joseph C. Chan\textsuperscript{a}, BS, Blackford Middleton\textsuperscript{a,b}, MD, MPH, MSc

\textsuperscript{a}Division of General Medicine, Brigham and Women’s Hospital, Boston, MA
\textsuperscript{b}Clinical Informatics Research and Development, Partners HealthCare, Boston, MA

Please address correspondence to:
Jeffrey A. Linder, MD, MPH
Division of General Medicine
Brigham and Women’s Hospital
1620 Tremont Street, BC-3-2X
Boston, MA 02120
Tel (617) 525-6654 / Fax (617) 732-7072
jlinder@partners.org

\textbf{Word Count:} 4,278
Abstract

For applications that require clinician use while interacting with patients, usability testing with standardized patients has the potential to approximate actual patient care in a controlled setting. We used hypothetical scenarios and a standardized patient to collect quantitative and qualitative results in testing an early prototype of a new application, the Acute Respiratory Infection (ARI) Smart Form. The standardized patient fit well into the usability testing sessions. Clinicians had a positive response to the standardized patients and behaved as they normally would during a clinical encounter. Positive findings of the ARI Smart Form included that users thought it had impressive functionality and the potential to save time. Criticism focused on the visual design, which could be streamlined, and navigation, which was difficult in some areas. Based on these results, we are modifying the ARI Smart Form in preparation for use in actual patient care. Standardized patients should be considered for usability testing, especially if an application is to be used during the patient interview.

Medical Subject Headings: Medical Record Systems, Computerized; User-Computer Interface; Evaluation Studies; Patient Simulation; Human Engineering; Respiratory Tract Infections

Non-MEsH Keywords: Human-Computer Interaction; Human Factors Engineering
1. Introduction

Electronic health records (EHRs) have tremendous potential to improve the quality of medical care by providing clinical decision support where and when it is needed. However, the full potential of EHRs has yet to be realized partially because of their poor uptake. One reason for the poor uptake of EHRs is lack of usability and integration into the clinical workflow. Currently, clinicians collect relevant information; process information; make clinical decisions; order appropriate tests, medications, or referrals; and then document each of these steps. One way to improve workflow would be by integrating all of these tasks into a single problem-oriented application.

We are developing problem-oriented “Smart Forms.” Web-enabled Smart Forms actively engage clinicians during the clinical documentation workflow, enhance the documentation process, and integrate decision support for several aspects of clinical decision-making, such as diagnostic test selection, medication prescribing, therapeutic planning, and patient education. One of the first Smart Forms we have designed is for acute respiratory infections (ARIs).

To improve quality of care, ARIs are a natural place to start. ARIs – including non-specific upper respiratory tract infections, otitis media, sinusitis, pharyngitis, acute bronchitis, influenza, and pneumonia – are the most common symptomatic reason for seeking care in the United States, accounting for 7% of all ambulatory visits [1]. In addition, ARIs are the number one reason for antibiotic prescribing in the United States; ARIs account for about 50% of all antibiotic prescriptions to adults [2]. However, much antibiotic prescribing for ARIs is inappropriate, which exposes patients to potential adverse drug events [3, 4], increases the prevalence of antibiotic-resistant bacteria [5], and increases costs. Use of an ARI Smart Form has the potential to improve clinician workflow and decrease inappropriate antibiotic prescribing [6].
A major challenge in designing usable applications for acute problems is that decision support is based on information collected during the visit and, to be actionable, must also be delivered at the time of the visit. By necessity, clinicians will need to interact with the ARI Smart Form while talking to patients. For such a system, standardized patients can play an important role in usability testing.

Usability has been defined as “the capacity to allow users to carry out their tasks safely, effectively, efficiently, and enjoyably” and is important in the design and improvement of electronic health record applications [7]. Usability testing generally involves a sample of target users performing tasks related to an application under development [8]. Usability testing can be performed under controlled experimental conditions in a “usability lab” or as naturalistic observations of clinicians in their usual work settings. As stimulus, usability engineers generally ask testing subjects to perform certain tasks or use written vignettes and to think aloud while they work with the system to complete these tasks. For applications that require clinician use while interacting with patients, usability testing with standardized patients has the potential to more closely approximate actual patient care in a controlled setting [9].

Standardized patients are “people trained to present, accurately and reproducibly, the problems that would be evident in a real patient” [10]. Standardized patients have been used for 30 years in medical education for teaching and evaluation [11], but more recently “unannounced” standardized patients have been used to evaluate the quality of administrative data [12] and actual physician performance [13]. Standardized patients have been generally found to be realistic [9] and able to portray the cases they have been trained to simulate [13, 14].

To evaluate the Acute Respiratory Infection (ARI) Smart Form, collect feedback about the effects of specific design decisions on user performance and satisfaction, and also provide
insight into user expectations, we performed usability testing using vignettes and a standardized patient.
2. Methods

2.1 The ARI Smart Form

The ARI Smart Form consists of 6 components: entry of clinical information; display of imported data; selection of diagnosis; provision of treatment options with integrated decision support; printing of patient handouts; and access to supporting medical literature (Figure 1). First, to support easy data capture, clinicians use radio buttons and drop-down menus to rapidly enter information about the details of the clinical work-up including chief complaint, history of present illness, symptoms, remedies tried, overall clinical course, review of systems, and physical examination. Second, the ARI Smart Form automatically imports the patient’s problem list, allergies, and medications. Third, the ARI Smart Form allows easy selection of the diagnosis using radio buttons for the primary diagnosis and checkboxes for any secondary diagnoses. Fourth, the ARI Smart Form provides diagnosis-specific treatment options with integrated clinical decision support, based on the “Principles of Appropriate Antibiotic Use for Treatment of Acute Respiratory Tract Infections in Adults” from the Centers for Disease Control and Prevention and the American College of Physicians [15]. For example, based on the presence or absence of certain signs and symptoms, the ARI Smart Form calculates the probability of a patient having streptococcal pharyngitis and recommends a course of action.

Fifth, for patient education, the ARI Smart Form provides one-click access to handouts regarding diagnoses, treatments, and “excuse-from-work” notes. Finally, there are links to the current guidelines from the Centers for Disease Control and Prevention, the American College of Physicians, and the Infectious Diseases Society of America. Once all information has been entered, the ARI Smart Form formats the information as a typical narrative note. The present prototype of the ARI Smart Form was designed to fulfill criteria for a “comprehensive” visit for
an established patient (Evaluation and Management Level 4) and take less than 15 minutes to complete.

In keeping with the fundamental purpose of Smart Forms, we have sought to integrate as many steps as possible. For example, by checking off a recommendation to order a medication, the medication list in the EHR is updated, the prescription is printed, and the assessment and plan section of the visit note documents the action. The ARI Smart Form runs within the Longitudinal Medical Record (LMR), the official ambulatory electronic health record of Partners HealthCare.

2.2 Usability Testing

The Institutional Review Board of Brigham and Women’s Hospital approved the study protocol. We conducted usability testing between January 24, 2005 and February 4, 2005, generally following the 9-step method of Kushniruk and Patel [7]. A group of 8 test participants – all physicians in the Partners provider network – worked with a prototype version of the ARI Smart Form [16]. The test participants had no prior exposure to the Smart Form prototype and were given only a basic introduction to the application in order to obtain their initial reactions to the user interface and enhance the quality of their feedback.

Test participants were given a set of 3 scenarios that involved using the ARI Smart Form prototype to perform specific tasks using hypothetical patient data. The 3 scenarios corresponded to 1) a 40 year-old with acute cough/acute bronchitis; 2) a 29 year-old preschool teacher with streptococcal sore throat; and 3) a 39 year-old with hypertension and hypercholesterolemia with a non-specific upper respiratory tract infection. The first 2 test scenarios were presented in written form (the first scenario was 6 sentences and the second scenario was 9 sentences). A standardized patient was used for the last scenario without written stimulus. We matter-of-factly
introduced the standardized patient to the test subjects and did nothing to suggest the use of standardized patients was atypical.

2.3 Standardized Patient

Many of the issues of standardized patients that arise in clinical performance assessment – such as reliability, discriminatory ability, and the recall of student performance by standardized patients [9, 14, 17] – are less important or not applicable to usability testing. Still some degree of realism is required. The standardized patient was trained with one of the investigators in an approximately 45 minute session. The standardized patient was educated about the classical findings of non-specific upper respiratory infections (e.g., runny nose, sore throat, cough, but no signs of more serious illness like fever, chills, or vomiting) and was given a basic “script” from which to work (the script was 9 sentences long). The standardized patient was instructed to try to directly answer questions, providing conversational answers, and succinctly improvise if the test participant asked questions that were not included in the script. The investigator and the standardized patient reviewed the scenario 3 times to practice and ensure consistency.

2.4 Subject Selection

We strove to examine a range of representative users performing representative tasks [7]. We requested participation from a variety of Partners HealthCare-affiliated primary care and urgent care sites via email. We asked initial respondents basic information about their use of the LMR. We selected our sample of 8 subjects, all attending physicians, stratified by the primary variable of LMR note-writing during patient visits. Because the ARI Smart Form was designed to be used during patient encounters, we were interested in getting a range of responses from test
participants who rarely, sometimes, or usually used the LMR to write notes while interviewing the patient. Test participants received a $50 gift certificate to an online bookstore for their participation in this study.

2.5 Setting

The test sessions were conducted in the offices of each participating physician using his or her own workstation (standard-build Pentium PCs, running Windows 2000 and Internet Explorer 6.0). During 4 of the 8 test sessions, the test administrator – an experienced usability engineer – and the standardized patient, who also served as an observer, were in the room with the participant. For the other 4 test sessions, the standardized patient served as the test administrator and there were no additional observers. The test sessions were scheduled to take 1 hour.

2.6 Measures

We collected quantitative and qualitative data for this study, as well as basic demographic information about the test participants and their usage patterns of the LMR. We asked test participants what tasks they perform using the LMR, what type of documentation style they prefer (dictated, free form, templated), and how many hours per week they use the LMR.

Quantitative outcome measures included the successful task completion rate and survey responses. Upon completion of each scenario, the test participants rated the software using a questionnaire that addressed their satisfaction regarding ease-of-use, time, and support information [18]. Test participants rated their level of agreement on a 7-point Likert scale ranging from Strongly Agree (1) to Strongly Disagree (7). Following the completion of all three
scenarios, the test participants filled out a post-test questionnaire that assessed their overall impressions of the ARI Smart Form [18]. One of the test participants used his entire session to provide feedback on just the first scenario. As a result, data was collected on the last two scenarios with only 7 participants.

For qualitative measures, test participants were encouraged to “think aloud” as they worked with the prototype so that the test administrator could more easily identify their attitudes and impressions while using the system [19]. Test participants were also encouraged to “think aloud” during the standardized patient scenario. At the end of each scenario and test session, the administrator and observer conducted a debriefing session with the test participant for a more detailed discussion of their experience with the application. Special attention was given to comments focusing on user interface issues and suggestions for improvement. The test sessions were audio recorded for later review. Video recording equipment was not used because of logistical issues including setup time, participant schedules, and the varying physical layouts of the test subject offices.

2.7 Data Analysis

Upon completion of the study, the test administrator and members of the research team reviewed the audio recordings, written notes, and design recommendations collected from each session. We constructed thematic frameworks using the method of Patton, who suggests reducing the volume of raw information by identifying patterns [20]. Using this data reduction method, core themes of participant feedback were then labeled using “sensitizing concepts” created by the authors to help make sense of the data to be presented. We present the primary themes and
augment them with context-specific user quotes collected from think-aloud and debriefing feedback.
3. Results

3.1 Characteristics of Test Participants

We selected the test participants stratified by electronic note-writing during patient visits: 3 rarely wrote notes during the visit, 3 sometimes, and 2 most of the time. Six of the 8 test participants were women and the age range was 37 to 63 years, with a mean age of 51 years. The test participants used the LMR for clinical documentation, tracking lab results, and prescribing medications. Two users preferred free form notes and 6 used structured templates. One of the test participants dictated notes infrequently. The amount of LMR use both during and outside of clinical sessions ranged from 5-10 hours per week to over 20 hours per week, with 4 of the test participants using the LMR over 20 hours per week.

3.2 Reaction to the Standardized Patient

There was no difficulty transitioning from written vignettes to the standardized patient from scenario 2 to scenario 3 and none of the test participants found the use of a standardized patient strange or difficult. Clinicians seemed to have a positive response to the standardized patient and were quite comfortable acting as they normally would during an actual encounter. Because they were instructed to think aloud, there were moments when some of the participants broke their exam protocol to elaborate on a user interface recommendation or a workflow task, but their behavior with the standardized patient was otherwise consistent with a typical visit.

3.3 Quantitative Results

Except for the test participant who spent the entire session on scenario 1, all of the test participants were able to complete the scenario tasks. Test participant responses to the post-
scenario questionnaires suggest an increase in their satisfaction with the ARI Smart Form as they progressed through the scenarios (Table 1). Five test participants commented that they found the ARI Smart Form prototype progressively easier and faster to use. The best scores were obtained with the final scenario, which featured the standardized patient.

A summary of the post-study questionnaire data (Table 2) reveals that the test participants were moderately satisfied with the ease-of-use of the ARI Smart Form prototype. The test participants rated their comfort and “ease of learning” the ARI Smart Form relatively higher than other questions. The ability to recover from mistakes, the pleasantness of the interface, and the organization of information in the Smart Form were rated relatively lower.

3.4 Qualitative Results

Qualitative results fell into five main themes. First, the test participants found navigation problematic in some places. The navigation links at the top of the form did not stand out and only one test participant identified them as navigational aids. The medical history list in the upper right-hand corner of the screen (Figure 1) allows users to view patient problems, allergies, and medications, but its layout and overall appearance were considered non-intuitive by nearly all users.

The second usability issue concerned the list of symptoms: test participants found it too detailed. Four of the users discussed how they normally use their own customized templates for specific ARIs, noting that they were more focused. Some test participants commented on the lack of a “symptom hierarchy.” For example, one participant said “If there is not cough, all cough-related questions should be grayed out.” Likewise, 4 test participants indicated that the ARI
Smart Form should be oriented to a “detailed” visit (Level 3), not a comprehensive visit (Level 4).

Third, test participants had problems with the visibility of recommendations. Due to the position of the “Recommendations” panel and the fact that nearly all of the test participants did not scroll all the way down, the buttons to print patient materials were not evident. One participant noted, “This is great, but it wasn’t intuitive to me that I could get that.” The title “Recommendations” itself was identified as being somewhat condescending and one test participant stated that he did not need to be given recommendations. “Plan of Action” was one of the suggested alternatives.

Fourth, test participants identified problems with the visual design of the ARI Smart Form. Most notable were the heavily saturated bright red and green background colors for controls in the symptoms table. Nearly all of the test participants found this color combination “too striking” and potentially non-intuitive (green for “normal” and red for “the presence of an abnormality”). In the Medical History section of the screen, there is a free text box displayed at the bottom of the window to enter smoking status. However, the grey background behind the text box field visually “separates” it from the “Smoking” label above it, making it unclear if the box is related to “Smoking” or the entire history list (problems, allergies, medications, smoking).

There were several typefaces used in the ARI Smart Form prototype, including serif and sans serif fonts, making the appearance somewhat cluttered. The “select all” check box at the top of the Symptoms table – intended as a shortcut to check the absence of all symptoms or abnormal physical findings – was not immediately obvious to any of the test participants. “It takes a long time to click all these radio buttons” was a frequent comment.
Fifth, test participants had comments about the format of the generated note. Some were concerned that although the note was automatically generated for them, they would still have to review it—a process that would take too much time. One participant commented, “My sense was that I had to edit the note pretty carefully and that I couldn’t just click on ‘Note’ and know that it was the way I want it.” It also was not immediately obvious to test participants they could edit the note once it was generated.

Additional comments from the test participants alluded to the lack of a save function to prevent loss of data in case of power or system outages; unclear abbreviations (e.g., “no MRG” for “no murmurs, rubs, or gallops” in the cardiac exam); the impersonal nature of using the computer while interviewing the patient; and concern about using the ARI Smart Form for patients with additional, non-ARI problems. It was also noted that some offices and examination rooms are not conducive to using the computer while talking to the patient.

Overall, the general feeling was that the ARI Smart Form might save time if users had the opportunity to grow accustomed to it, but 2 participants explicitly stated they would probably not use it and instead would continue using their own templates.
4. Discussion

4.1 Use of a Standardized Patient

In usability testing of a novel EHR application for an acute problem, we used a standardized patient, which was well accepted by test subjects and produced useful results. Others have described the use of a usability expert playing the role of a patient [21], but we would argue that in certain situations it may be more helpful to have a dedicated, standardized patient provide as the stimulus to usability testing subjects. Use of a dedicated, standardized patient will ensure consistency in the case presentation, the realism of the test situation, and enhance the ability of the test facilitator to more accurately record their observations of the test subject.

There are certain situations in which standardized patients will be particularly helpful. First, and most obviously, standardized patients should be considered in situations where the application requires both the patient and clinician. Standardized patients will not be useful for applications that are not dependent on the presence of both the patient, such as a “results manager” [8], or the physician, such as home-based applications [22]. Second, standardized patients can be helpful for those applications for which timing is critical. For ARI visits that are typically very brief, it is important to get test subjects’ perceptions of time as they are talking to the standardized patient. Third, standardized patients should be considered when the interaction of the clinician and the patient is integral to the application, such as applications that require gathering a history. Fourth, special attention should be paid to situations where an application has the potential to interfere with the patient-physician interaction. Though computer use can have favorable effects on the patient-physician interaction [23], applications that are hard to use and draw attention away from the patient are likely to be rejected by clinicians.
4.2 Summary of Usability Results

Based on both quantitative and qualitative results from the vignettes and the standardized patient, the overall attitude towards the ARI Smart Form was positive. Test participants grew progressively more comfortable and satisfied with the ARI Smart Form as they used it. This is typical as users gain experience with a system and apply what they have learned to later scenarios. Two test participants stated they would probably continue using their own templates. Future versions of the ARI Smart Form must provide such clinicians clear, immediate value if the ARI Smart Form is to be widely adopted.

4.3 Implications for Future Versions

We identified a number of potential modifications to enhance the usability, acceptability, and value of the ARI Smart Form [24]. The quantitative results indicate that our improvement efforts should focus on clarifying how to recover from mistakes, the quantity and organization of information, and the quality of the interface.

The qualitative results brought up some concrete improvements to be made to the ARI Smart Form, such as the color scheme and facilitating navigation [25]. We have oriented the ARI Smart Form more towards a “detailed” (Level 3) visit, clarified the handling of smoking status, and changed the title of the “Recommendations” to “Orders/Assessment/Plan.” We have also improved the appearance of the generated note and now allow users to move back and forth between the template and the generated note.

Test participants also identified some challenging areas. While most test participants felt the list of symptoms was too long, we are caught between providing a much shorter list and
maintaining a list that is applicable to all ARI diagnoses [8]. To shorten the list, future prototypes will use symptom dependencies. For example, as suggested, users will need to select “cough” before they can select “productive cough.”

Test participants’ difficulty with the “all-normal” checkbox is an excellent example of how design problems can lead to workflow inefficiencies. We designed the “all-normal” checkbox to be a major time-saver; clinicians can document the normality of all symptoms and physical exam findings in just one click. However, this feature was not obvious to nearly all of the test participants. In future prototypes, we will show a label for the “all-normal” checkbox when the user “mouses over” it with the pointer (Figure 2).

4.4 Limitations

This study has limitations that should be considered. First, we used a small number of physicians. Although a larger number of test participants is generally unnecessary to identify additional usability concerns [7, 16] and we tried to draw a representative sample of test subjects, subjects in this study may or may not be representative of primary care clinicians in general. We included clinicians with a range of rates of computer use as they interviewed patients. Second, the ARI Smart Form was in a prototype stage and lacked some functionality we anticipate in the final version. Third, we were actively working out “bugs,” such as the problem, allergy, and medication headers in Medical History that were not “active” at the time of usability testing. Fourth, we provided our test participants with relatively little training prior to use. Actual clinical users will have a formal introduction and have access to online help if they encounter problems. Fifth, we did not formally evaluate participants’ reaction to the standardized patient or statistically compare the data obtained in the presence or absence of the standardized patient
because of the small sample size. The matter-of-fact presentation of the standardized patient as a
stimulus may have contributed to the acceptability of the standardized patient. Scenario 3, with
the standardized patient, received better scores than the other 2 scenarios, but we are unable to
tell if this was due to the standardized patient or simply due to users’ increasing familiarity with
the application. Moreover, the purpose of a standardized patient was not to increase user
satisfaction but to increase the quality and quantity of information regarding usability. To that
end, it may be enough to know that the standardized patient fit well into the usability testing
sessions and we obtained useful information regarding the user interface. Some of these
limitations could be addressed with a larger, randomized controlled trial of standardized patient
use on usability outcomes, reliability of results, and test participants’ impressions.

Although this usability testing was done with an application for our locally developed
EHR, the insights gained are generalizable to usability testing for other applications and other
EHRs. To improve quality of care, researchers, health policy experts, clinical leaders, and
developers are struggling to broaden the reach of clinical decision support technology [26].
Clinical decision support and EHRs will not reach their full potential if they are not maximally
usable, integrated into the clinical workflow, and do not provide obvious value to clinicians.
Smart Forms, by integrating the provision of clinical information, ease-of-documentation,
decision support, patient education, and ordering, have the potential to enhance clinician
workflow, increase the use of problem-oriented clinical decision support and, more generally,
increase the use of EHRs, if they are well designed.
5. Conclusions

We conducted usability testing using a standardized patient for a novel application that requires clinician-patient interaction. The standardized patient fit well into the usability testing sessions and was helpful in generating useful data. The findings of the present study are being used to improve our application. Standardized patients should be considered for usability testing, especially if an application is to be used during the patient interview, has the potential to interfere with patient-clinician communication, or if timing is critical.
Acknowledgement

Supported in part by grants from the Agency for HealthCare Research and Quality (HS015169, HS014563, and HS014420) and the National Heart, Lung, and Blood Institute (HL072806).
References


Table 1: Post-Scenario Questionnaire Results

Table 2: Post-Study Questionnaire Results

Figure 1: Initial Prototype of the ARI Smart Form

Figure 2: Screen Shot of Redesigned ARI Smart Form

* Scored from 1, strongly agree, to 7, strongly disagree. The first 2 scenarios used written vignettes. The third scenario used a standardized patient.

† Scored from 1, strongly agree, to 7, strongly disagree.