The Remote Patient Education in a Telemedicine Environment Architecture (REPETE)

Albert M. Lai, Ph.D.,[1] Justin B. Starren, M.D., Ph.D.,[2] David R. Kaufman, Ph.D.,[1,3] Eneida A. Mendonça, M.D., Ph.D.,[1] Walter Palmas, M.D.,[4] Jason Nieh, Ph.D.,[5] and Steven Shea, M.D., M.S.,[1,4,6] for the IDEATel Consortium[7]

Departments of [1]Biomedical Informatics, [3]Psychiatry, [4]Medicine, [5]Computer Science, and [6]Epidemiology, Columbia University, New York, New York.
[2]Biomedical Informatics Research Center, Marshfield Clinic, Marshfield, Wisconsin.
[7]http://www.ideatel.org/acknowledgement.html

Abstract

The objective of the study was to develop and implement an architecture for remote training that can be used in the narrowband home telemedicine environment. A remote training architecture, the REmote Patient Education in a Telemedicine Environment (REPETE) architecture, using a remote control protocol (RCP) was developed. A set of design criteria was specified. The developed architecture was integrated into the IDEATel home telemedicine unit (HTU) and evaluated against these design criteria using a combination of technical and expert evaluations. Technical evaluation of the architecture demonstrated that remote cursor movements and positioning displayed on the HTU were smooth and effectively real-time. The trainers were able to observe within approximately 2 seconds lag what the patient sees on their HTU screen. Evaluation of the architecture by experts was favorable. Responses to a Likert scale questionnaire regarding audio quality and remote control performance indicated that the expert evaluators thought that the audio quality and remote control performance were adequate for remote training. All evaluators strongly agreed that the system would be useful for training patients. The REPETE architecture supports basic training needs over a narrowband dial-up connection. We were able to maintain an audio chat simultaneously with performing a remote training session, while maintaining both acceptable audio quality and remote control performance. The RCP provides a mechanism to provide training without requiring a trainer to go to the patient's home and effectively supports deictic referencing to on screen objects.

Key words: *remote training architecture, REPETE, home telemedicine unit*

Introduction

Despite best efforts to make home healthcare technology easy to use, patients frequently require multiple training sessions to master the skills needed to use these devices, particularly devices that include Web access. In geographically distributed telemedicine projects such as IDEATel,[1] where the use of telemedicine can be most advantageous, repetitive traditional training sessions can be costly because of the amount of time that is necessary to reach the patients' homes.[2]

Conventional telephone support solutions are frequently unsatisfactory. Elderly telemedicine patients frequently lack the expressive vocabulary for describing components of graphical user interfaces such as scroll bars and other widgets.[3,4] This is compounded by problems associated with older adults' lack of visual acuity and limits in their ability to selectively attend to relevant screen features.[5] Therefore, even if the trainer can precisely describe to the patient what they should be looking for on their Home Telemedicine Unit (HTU), the patient may not be able to connect the trainer's words to the images on the screen. A potentially viable solution to this problem would be to develop a method to remotely educate patients to use their home healthcare devices. Remote training may be more cost effective, and can be used to provide more frequent training. In addition, it may be possible to

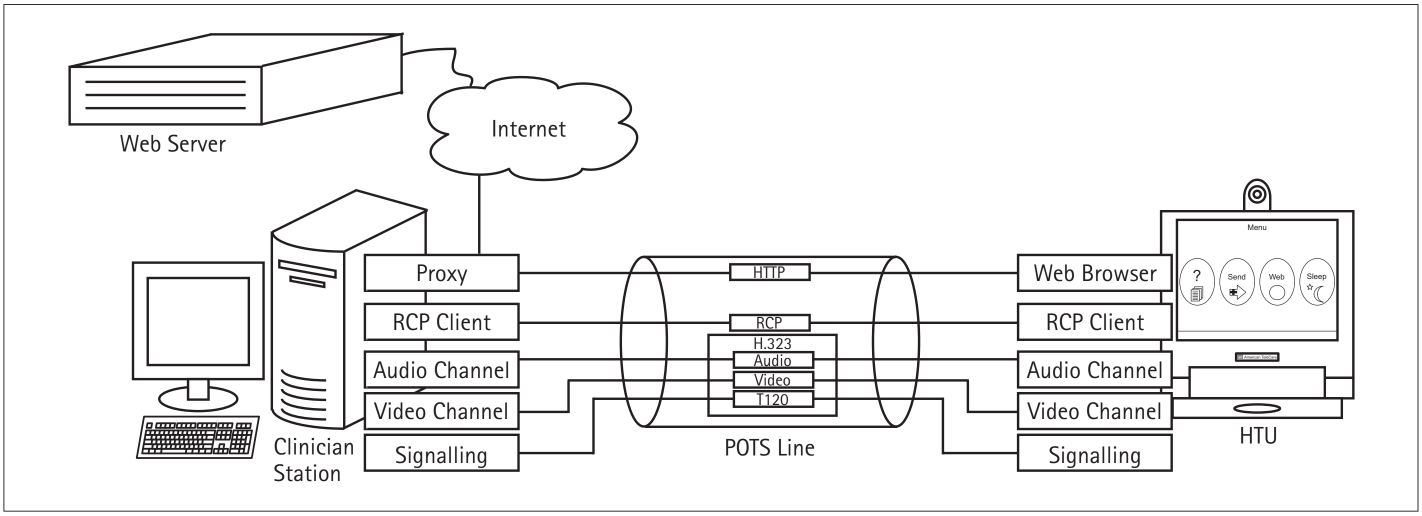

Fig. 1. REPETE architecture.

provide assistance on demand to patients, rather than for them to wait for a trainer to travel to patients' homes.

- To address these difficulties, the REmote Patient Education in a Telemedicine Environment (REPETE) architecture was developed. This architecture provides a mechanism to provide HTU training to home telemedicine patients without requiring a trainer to travel to the patient's home.

This paper discusses the design criteria for a remote training architecture; perceptual and cognitive issues for remote training; the architecture that was developed; the optimization methodology used; and an evaluation of the technical performance of the architecture.

Material and Methods

DESIGN CRITERIA

A task analysis of the interaction between the trainer and patient was performed. A typical scenario would be for the trainer to demonstrate or guide the patient to perform an action. To do this, the trainer must be able to move the cursor and to engage in clear voice communication. After guidance and demonstration by the trainer, the patient would then perform the requested action, often involving interaction with the project Web portal or other Web site. This analysis led to the following design criteria:

- The system needed to accommodate the entire training environment over a single telephone line.
- The system needed to support audio chat so that trainer and patient could converse during the session.
- The system needed to support browsing of external Web sites during the session so that issues of content interpretation could be assessed.
- To the extent possible, the system should utilize open-source solutions.
- The system should be able to control all portions of the HTU software to support remote maintenance of HTU settings.
- The system needed to be implemented using existing telemedicine hardware employed in the IDEATel project.

ARCHITECTURE DESIGN

The REPETE architecture (*Fig. 1*) leverages existing H.323 videoconferencing infrastructure used by many home telehealth devices. In addition to the H.323 capabilities, remote training is supported through the use of a remote control protocol (RCP). This architecture allows for simultaneous voice conferencing and remote control over a single telephone line, providing the trainer and patient shared control over the HTU's workspace.

We selected Virtual Network Computing (VNC) as our RCP.[6] Among the VNC implementations evaluated, UltraVNC (UVNC B.V.B.A., Antwerpen, Belgium) was selected because it offered several useful features, most importantly server-side scaling. To support Web browsing from the HTU, the provider station acts as a Web-proxy server. The HTU then connects to the provider/proxy, which then forwards HTTP requests to the Internet over its existing broadband connection.

We modified American TeleCare's (ATI) (American Telecare, Inc., Eden Prairie, MN) NX HTU and provider stations as used by the

IDEATel project to include REPETE. A code-sharing agreement with ATI to make the changes necessary to create a prototype implementation for remote training on their HTUs was reached. ATI subsequently implemented these changes into the production telemedicine system. A screenshot of the trainer's workspace is shown in *Figure 2*.

OPTIMIZATION

Because the performance of the RCP over a modem is bandwidth bound, a tradeoff between speed and screen fidelity is possible. Although research in this area is limited, studies suggest that the bandwidth of screen updates can be significantly reduced by reducing screen fidelity through downsampling, especially when coupled with a reduction of bit depth.[7,8] Although the actual resolution of the HTU display was 800 × 600 at 24-bit depth, the transmitted image could be downsampled by a factor of 2 or 4 (400 × 300 or 200 × 150) and the bit depth could be reduced. On the trainer's workstation, the downsampled image is upscaled back to the original resolution. With high levels of downsampling, some objects on the screen such as text could become illegible because of the loss of resolution. Initial tests revealed that fourfold downsampling and bit depths less than 8 produced unacceptable images, and therefore, were not evaluated further. Bit depths of 24 and 8, with and without twofold downsampling, however, appeared to be acceptable and were further evaluated using the methodology described below.

OPTIMIZATION METHODOLOGY

An experimental testbed was set up to determine what VNC settings could be used feasibly. The measurement methodology used was based on a technique called slow-motion benchmarking.[9] Packet capture software, Ethereal 0.99.0 (Ethereal, Inc., Lenexa, KS) with WinPcap 3.1 (CACE Technologies, Davis, CA), was run on the trainer's workstation. The provider station and the HTU were connected to each other via a Teltone TLS-3B (Teltone Corporation, Bothell, WA) telephone line simulator (*Fig. 3*). The phone-line simulator allowed the prototype provider station and HTU to dial each other without needing a real telephone line. Trainer-side network activity was used as a means of measuring the latency of a screen update from user input to the time the screen update arrives on the trainer's workstation.

In the experimental scenario, the HTU ran the VNC software continuously in the background. The provider station dialed the HTU, similarly to what would occur during a normal televisit. Instead of proceeding to a video visit immediately, the provider station sent a command to the HTU that caused the HTU to go to its home screen. From the provider station, a VNC connection was established to the HTU. Using this VNC connection, the HTU was controlled from the provider station, first launching the Web browser. The Web browser was then directed to a patient login screen. The username and password for the Web site were entered into their respective text boxes. In order to prevent any extraneous VNC network traffic, the cursor was clicked outside of the text boxes to prevent the cursor from blinking. Ethereal was then launched on the provider station. Packet capture was started and the network activity was monitored from the time the login button was pressed from the trainer station until the myIDEATel summary page was fully loaded. Visual comparisons of cursor positioning and movements generated on the trainer's workstation to those transmitted to and displayed on the HTU were also noted. Observations regarding dynamic positioning of the cursor on the trainer's workstation compared to that on the HTU were made. In addition, smoothness of the cursor movement as it was translated onto the HTU was also noted. Music was played continuously across the H.323 audio channel in order to gauge the level of audio dropout during page downloads. A simple subjective approach to

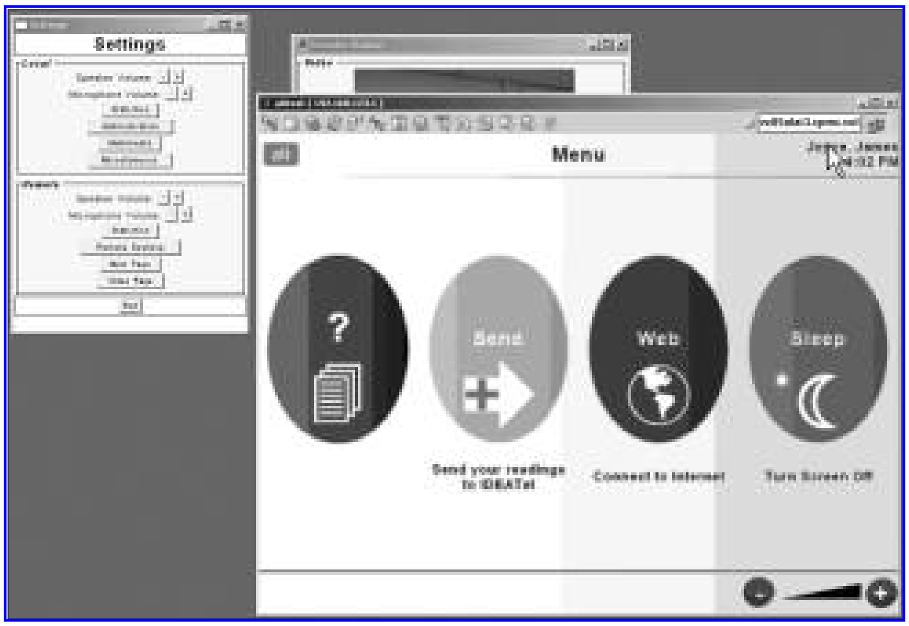

Fig. 2. Screenshot of modification of ATI's NX home telemedicine unit including REPETE, showing trainer's workspace.

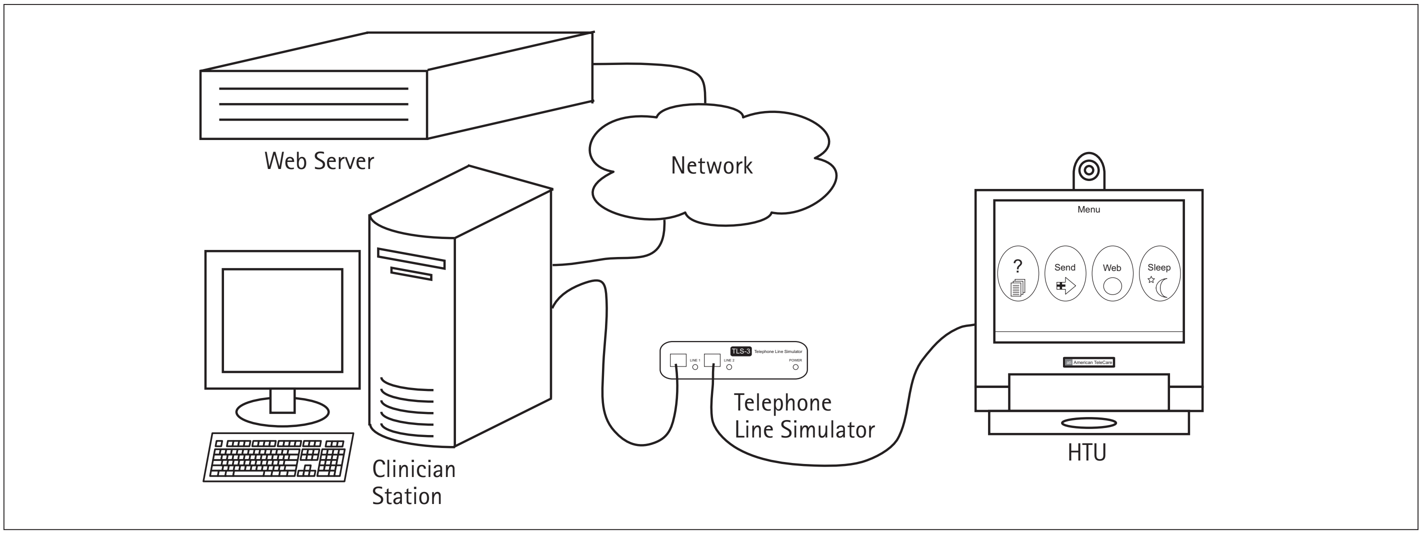

Fig. 3. Experimental testbed showing trainer's workstation, Web server, telephone line simulator, connected to home telemedicine unit (HTU).

audio quality was used. The main subjective measures were whether or not the audio would be intelligible and the severity of audio dropout, if any, during the experiments.

This process was completed for each combination of VNC settings. Total page download time and trainer lag were measured using the packet capture. Total page download time measures the amount of time from the click of the mouse on the login screen until the myIDEATel summary screen is fully loaded and displayed on the trainer's workstation. The delay between the patient's screen view and the trainer's view (trainer lag) was determined by inspecting the packet traces.

Total page download time and trainer lag were computed using the following method. By inspecting the packet traces, one can see the packet generated by the trainer's workstation signaling the click of the mouse. Data are then transmitted from the HTU back to the trainer's workstation until the myIDEATel summary screen is fully loaded. When the screen update is fully loaded on the trainer's workstation, there are no longer any data being transmitted over the dial-up connection and packet capture is stopped. From this packet trace, the total page download time is determined by comparing the timestamp of the first packet generated by the click of the mouse and the last packet captured for the screen update. Similarly, the trainer lag is determined by finding the timestamp of the last HTTP packet transmitted to the HTU and comparing it to the timestamp of the last packet sent to the trainer's workstation for the screen update.

EXPERT EVALUATION

In addition to quantitative evaluation of REPETE, a series of user studies were performed. The architecture was evaluated by conducting mock training sessions on the IDEATel data review Web site, over the testbed, using the same set of VNC settings.

The evaluation was conducted by three members of the IDEATel project research staff, including a case manager, a trainer, and a researcher. Because the evaluation was conducted by project staff as part of the system development process, separate Institutional Review Board approval was not required.

Two core myIDEATel tasks were used as demonstration tasks (1) logging into the myIDEATel Web site and (2) reviewing monitoring data. Throughout the mock training sessions, the trainer spoke to the evaluator through the H.323 Voice Over Internet Protocol audio chat. The training sessions were guided using the RCP. After seeing how a training session could be performed, the evaluator was then guided through using the system as a trainer. After experiencing the remote training from both perspectives, the evaluators completed a 6-item questionnaire scored on a 5-point Likert scale. The questionnaire included questions on audio quality, the interactivity of remote control, and the remote control video quality. The questionnaire also included questions on the usefulness of the system. The full questionnaire is included in *Table 1*.

Results

OPTIMIZATION RESULTS

Figure 4 shows measurements taken from the packet captures

Table 1. Evaluation Questions

QUESTION	S1	S2	S3	MEAN	STD DEV.
For the clinician, the audio is adequate.	5	4	3.5	4.2	0.76
For the patient, the audio quality is adequate.	5	4	5	4.7	0.58
The video response for the remote control is adequate	4	4	4	4	0
The video quality for the remote control is adequate.	4	4	4.5	4.2	0.29
The system will be useful for training patients.	5	5	5	5	0
By using this method of home training, the number of home visits will be reduced.	5	3	3	3.7	1.15

with varying the bit depths and downsampling factors. Changes in the VNC settings significantly improved the response times for all three of the parameters evaluated: (1) total page download time, (2) trainer lag, and (3) Web request time. The amount of time for the HTU to download the myIDEATel summary Web page varied between 13 seconds and nearly 20 seconds. There was nearly a 100% improvement in total page download time between 24 bpp non-downsampled and the 8 bpp downsampled percent. The amount of lag that the trainer experiences in screen refreshes after the patient's HTU has loaded the Web page is nearly a 500% improvement between the 24 bpp non-downsampled and the 8 bpp downsampled settings.

The cursor movements and positioning displayed on the HTU were smooth and were effectively in real-time. Any delays were too small to be visually noticeable. All settings resulted in some audio dropout. The subjective difference between the level of audio dropout at 8 bpp non-downsampled and 8 bpp downsampled was marginal. However, at 24 bpp, both non-downsampled and downsampled resulted in larger durations where the audio quality was impacted than at 8 bbp because of the longer page download and screen update times. As expected, audio performance was impacted most during Web page downloads and screen refreshes. Audio was not noticeably affected, if at all, during mouse cursor movements.

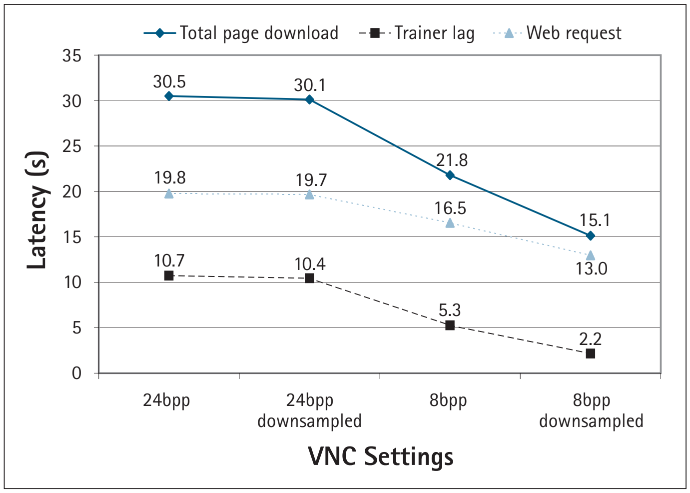

Fig. 4. Measurements taken from packet captures, varying the bit depths and downsampling factors.

EXPERT EVALUATION RESULTS

The overall response was favorable. Individual responses ranged from 3 to 5, where 1 is strongly disagree and 5 is strongly agree. Averages for questions ranged from 5 to 3.7. The average response for overall utility of the system was 5. When the experts were asked whether the system would reduce the need for home visits, the average response was 3.7 (range 3–5). The remaining questions on audio quality and remote control performance averaged between 4 and 4.6. Expert responses are shown in *Table 1*.

Discussion

The REPETE architecture was able to achieve all of the design criteria. It also provides more support for the old maxim that "the devil is in the details." Optimizing VNC for the dial-up network resulted in dramatic improvements in both quantitative performance and subjective user experience. There was nearly a fivefold reduction in the amount of screen update lag a trainer would experience.

There are few formalized studies that have studied the use of RCPs for training.[10] Previous studies have indicated that server-side screen scaling may be useful for providing thin client Web browsing access on small screen devices such as personal digital assistants over broadband network connections,[7,11] but the use of server-side scaling has not been studied in the context of training. This study was used as an initial feasibility study to determine whether or not an RCP could be used over a dial-up telephone line to provide remote training.

This study focused on the worst-case scenario—full screen updates. Smaller screen updates, such as drop-down menus and typing in text boxes, have much less lag time than full screen updates and may have smaller relative improvements than those reported here. These smaller updates are not likely to be bandwidth bound even in the low bandwidth environment of dial-up networking.

Some results were unexpected. In particular, at 24 bpp, the improvement in performance between the non-downsampled and the downsampled settings for all three measurements was small. However, the improvements between the non-downsampled and the downsampled settings for 8 bpp were very large. In addition, the improvement for the total page download time between 24 bpp downsampled to 8 bpp non-downsampled was marked, even though the theoretical amount of data was similar.

There was a dramatic improvement in performance for trainer lag times when moving from 24 bpp to 8 bpp and 8 bpp downsampled. The nearly 500% improvement enables the trainer to observe the screen on the HTU with far more accuracy. A 10-second lag between what the patient sees and what the trainer sees would result in a significant obstacle for the trainer to describe to the patient what is being shown on the screen or in guiding the patient to make various actions on the system. However, the approximately 2-second lag is comparatively negligible, in effect, allowing the trainer to seamlessly view what appears on the patient's HTU.

Audio performance was noticeably impacted during Web page downloads and screen refreshes. However, in the remote training scenario, it is anticipated that the audio dropout, particularly for the lower amounts of time exhibited in the lower bandwidth conditions of 8 bpp downsampled, would not be a significant problem. This is due to the fact that during the Web page downloads, not much dialogue between the patient and the trainer is expected to take place as both are waiting for system response and a screen refresh before they can continue with the training session. In contrast, audio performance was not affected noticeably during mouse cursor movements. This allows the trainer to use the mouse cursor as a telepointer for deictic referencing (i.e., establishing a common frame of reference for referring to "this" object or "that" object) and gesturing during a training session. In addition, these dropouts were not noticed by the expert evaluators. However, it remains to be seen whether they would have any discernible impact in a real-world training situation.

The evaluation by the expert evaluators confirmed our intuition and show that the REPETE architecture supports basic training needs over a narrowband dial-up connection. We were able to maintain an H.323 audio chat simultaneously with performing a remote training session, while maintaining both acceptable audio quality and remote control performance. In addition, there were no complaints over the reduced screen fidelity due to bit-depth reduction and server-side (HTU) screen scaling. Although the evaluators were unsure whether or not the availability of REPETE for remote training would reduce home visits, the evaluators overwhelmingly thought that the system would be useful in patient training.

REPETE was integrated into the existing IDEATel telemedicine architecture, and it is reasonable to hypothesize that it can be integrated into most advanced HTUs. Based on the analyses and testing performed, we have shown that an RCP has the potential to be used for remote training of HTUs in the dial-up networking environment of home telemedicine. An RCP provides a mechanism to provide training without requiring a trainer to go to the patient's home. This type of training scenario provides for one-on-one training, enabling the training to be self-paced and allows for the opportunity for real-time feedback. An RCP also effectively supports deictic referencing to on-screen objects through movement of the cursor and is likely to be effective in visual attention capture. The performance of REPETE over a dial-up connection was judged by experts to be acceptable through both quantitative and qualitative metrics. However, this is just a provisional test, and we are currently completing studies that assess the efficacy of this method in training older adults to use a telemedicine system.

Outside of the telemedicine environment, the REPETE architecture could be adapted to remote training in a variety of environments. The combination of H.323-audio and RCP could be implemented on most personal computers. This would allow training to occur in locations that are impractical to send trainers to, including rural locations and private homes. Many seniors who are not in telemedicine programs could also benefit from this sort of training for general Web applications.

Acknowledgments

This work was supported in part by Centers for Medicare and Medicaid Services Cooperative Agreement 95-C-90998 and National Library of Medicine Training Grant NO1-LM07079. We also acknowledge Jenia Pevzner and Iris Fernandez for their assistance and expertise.

REFERENCES

1. Starren J, Hripcsak G, Sengupta S, Abbruscato CR, Knudson PE, Weinstock RS, Shea S. Columbia University's Informatics for Diabetes Education and Telemedicine (IDEATel) project: Technical implementation. *J Am Med Inform Assoc* **2002**;9:25–36.

2. Starren J, Tsai C, Bakken S, Aidala A, Morin P, Hilliman C, Weinstock R, Goland R, Teresi J, Shea S. The role of nurses in installing telehealth technology in the home. *Comput Inform Nurs* **2005**;23:181–189.

3. Kaufman DR, Patel VL, Hilliman C, Morin PC, Pevzner J, Weinstock RS, Goland R, Shea S, Starren J. Usability in the real world: Assessing medical information technologies in patients' homes. *J Biomed Inform* **2003**;36:45–60.

4. Kaufman DR, Pevzner J, Hilliman C, Weinstock RS, Teresi J, Shea S, Starren J. Redesigning a telehealth diabetes management program for a digital divide seniors population. *Home Health Care Manage Pract* **2006**;18:223–234.

5. Fisk AD, Rogers WA, Charness N, Czaja SJ, Sharit J. *Designing for older adults: Principles and creative human factors approaches*. Boca Raton, FL: CRC Press, **2004**.

6. Richardson T, Stafford-Fraser Q, Wood KR, Hopper A. Virtual network computing. *IEEE Internet Comput* **1998**;2:33–38.

7. Kim J, Baratto R, Nieh J. pTHINC: A thin-client architecture for mobile wireless web. In: *Proceedings of the 15th World Wide Web Conference*, May 23–26, **2006**, Edinburgh, Scotland, pp.143–152.

8. Lai AM, Nieh J. On the performance of wide-area thin-client computing. *ACM Trans Computer Syst* **2006**;24:175–209.

9. Nieh J, Yang SJ, Novik N. Measuring thin-client performance using slow-motion benchmarking. *ACM Trans Computer Syst* **2003**;21:87–115.

10. Adams J, Rogers B, Hayne S, Mark G, Nash J, Leifer L. The effect of a telepointer on student performance and preference. *Computers Educ* **2005**;44:35–51.

11. Kim J, Baratto R, Nieh J. An application streaming service for mobile hand-held devices. Proceedings of the IEEE International Conference on Services Computing, September 18–22, **2006**, Chicago, IL, pp. 323–326.

Address reprint requests to:
Albert M. Lai, Ph.D.
Department of Biomedical Informatics
Columbia University
622 West 168th Street
Vanderbilt Clinic, 5th Floor
New York, NY 10032

E-mail: albert.lai@dbmi.columbia.edu

Received: July 24, 2007
Accepted: September 27, 2007