Journal of Biomedical Informatics 43 (2010) 75–80

Contents lists available at ScienceDirect

Journal of Biomedical Informatics

journal homepage: www.elsevier.com/locate/yjbin

# Development of a hand-held computer platform for real-time behavioral assessment of physicians and nurses

Nabyl Tejani [a], Timothy R. Dresselhaus [b,c], Matthew B. Weinger [d,e,*]

[a] Yale University, 300 George Street, New Haven, CT 06511, USA
[b] San Diego School of Medicine, The San Diego VA Health Care System, The San Diego Children's Hospital, USA
[c] Department of Medicine, University of California, San Diego, USA
[d] Departments of Anesthesiology, Biomedical Informatics, and Medical Education, Center for Perioperative Research in Quality, Vanderbilt University School of Medicine, Nashville, USA
[e] Geriatric Research Education and Clinical Center, Tennessee Valley VA Healthcare System, Nashville, TN, USA

ARTICLE INFO

*Article history:*
Received 3 December 2008
Available online 22 August 2009

*Keywords:*
User-centered design
Usability engineering
Ecological momentary assessment
Computer-assisted data collection
Response burden
Hand-held computer
Software development
Hospital-based clinicians

ABSTRACT

We developed a hand-held data collection tool to facilitate real-time collection of data on the factors that affect hospital staff performance. To assure high-yield of data from busy clinicians, the design objectives included low response burden, the ability to collect complex real-time data in dynamic work environments, and automated data integration. Iterative user-centered design of custom interfaces resulted in a dynamic intuitive platform where branching logic was applied to present a series of survey questions dependent on the participant's responses. Over a 12-month period, 304 inpatient physicians and nurses completed (with minimal initial training) a total of 11,381 survey responses. For randomly timed repeated survey prompts, complete (73%) or partial (12%) responses were obtained in a median time of 96 s.



## 1. Introduction

Improvements in health care quality require a thorough understanding of clinical processes of care. Hand-held computing technology can facilitate the efficient collection of real-time, complex information from individuals in the midst of dynamic work environments. However, there are significant challenges to creating and deploying an effective hand-held data collection tool that is minimally disruptive to clinicians with multiple higher priority patient care demands. We describe the development of custom survey software designed to randomly prompt hospital-based clinicians to complete surveys regarding their current clinical tasks as well as their mood, cognition and work demands. The challenges overcome and lessons learned have general applicability to the development of software tools for similar clinical research applications.

## 2. Background

Alternative strategies used to capture information about clinician work include retrospective logs and direct observation. While logs can be useful in less time intensive settings, they tend to be incomplete, inaccurate, and affected by recall bias, particularly when workload increases. Direct observation can be resource intensive in certain settings. Moreover, observers must balance interpreting participants' actions, thoughts, and intents versus asking the participants directly and thereby influencing the native environment. Ecological momentary assessment (EMA) is an alternative method for efficiently capturing self-reported data which has been shown to work well in the clinical setting. EMA uses intermittent, random sampling to generate information about work or other real-time variables. EMA has been used to assess patient health outcomes, including ambulatory blood pressure associations with negative emotions [1,2], mood and activity predictors of substance use [3,4], and emotional triggers of binge eating behavior [5,6]. EMA has become a popular method for understanding patients' behavior and their reaction to disease or treatment [7–9], but few studies have used EMA to study clinician practice or behavior [10]. We sought to use EMA in a multi-site evaluation of physicians and nurses to assess the relationship between clinical task demands, clinician factors (e.g., experience, mood, and cognition) and the occurrence of medication errors. This required a flexible, low-cost hand-held platform that would randomly prompt busy clinicians to respond, efficiently capture multidimensional survey responses, and facilitate data aggregation and analysis. We describe the development steps, functional

* Corresponding author. Fax: +1 615 936 7373.
*E-mail address:* matt.weinger@vanderbilt.edu (M. Weinger).

1532-0464/$ - see front matter © 2009 Elsevier Inc. All rights reserved.
doi:10.1016/j.jbi.2009.08.011

requirements and ultimate usability of this new health services research tool.

The hand-held data collection device was developed as part of a broader study on the nature of medication errors in inpatient adult and pediatric medical settings of four academic hospitals. The design of the full study was to link medication errors to intrinsic (clinician) and extrinsic factors. Medication errors were captured through well established systems of self-report of occurrences and pharmacist intervention at each of the participating institutions. For example, at one institution medication errors were captured and recorded in three databases: (1) unpreventable adverse drug reactions (ADR); (2) intercepted prescribing errors (interventions); (3) all other medication errors, including preventable adverse drug events (PADE). Intravenous pump dose error reduction software (the ALARIS Guardrails Software System for infusion devices) was also used to capture medication programming dose errors.

## 3. Design principles and objectives

The design requirements were prospectively identified based on observations of and input from physicians and nurses, as well as on prior experience [10] and appreciation of principles of usability engineering and design [11–13]. These principles include the following high-level design considerations intended to help designers produce fundamentally correct user interfaces [14]:

1. *Seek user input early and often*.
2. *Focus on user performance, not preference*.
3. *Err on the side of design simplicity*.
4. *Anticipate device failure – make it more obvious and help users recover*.
5. *Facilitate work flow – Avoid negative effects of device use on task flow*.
6. *Enable users to set the pace*.

The design requirements were identified as follows:

1. *Accurate and reliable data collection*. The accuracy of the self-entered data should be easily verifiable by the participant. Data entered should be reliably preserved (up to 1 week) until downloaded to a master database.
2. *Intermittent random sampling*. Sampling should occur randomly within prescribed time intervals (e.g., 90 min) for the intended duration of the survey period (1 week).
3. *Intuitive, easy-to-use interface*. Questions should be presented in a manner that allows data entry by participants under time pressure who have no prior experience using the instrument. Many participants were not familiar with the Palm hand-held computer or Graffiti™.
4. *Low response burden*. Participants were to be sampled repeatedly (i.e., 25–50 times) during busy clinical work. The design goal was less than 2 min per random survey response.
5. *Security and confidentiality*. Both participants and the Institutional Review Boards expected that confidential information would be de-identified and kept secure.
6. *Adaptability*. The survey application should be easily modifiable to present different research questions to different end-user groups (e.g., to present different questions to a clinical pharmacist than to a physician). The software also had to be compatible with multiple hand-held computer platforms.
7. *Ease of data processing integration*. Data should be downloaded and consolidated in an efficient and automated manner from multiple devices. The tool should provide for efficient transfer of participant responses into standard data analysis software running on standard desktop hardware.
8. *Low cost*. To allow deployment of a large number of hand-held devices, cost had to be minimized.

## 4. System description and design process

### 4.1. Tool development

The selection of a development platform presented a challenge due to the many options available, each with strengths and limitations. The platforms evaluated included CodeWarrior, Falch.net, CASL, and Appforge. Appforge was ultimately selected because (1) we could write code in Visual Basic which was intuitive and easily modified; (2) we could set multiple alarms at different times as well as multiple timers of different durations; (3) it included a customizable conduit to Hotsync data between the hand-helds and a Microsoft Access database residing on a desktop PC; (4) it could easily convert our application to run on multiple platforms. Offsetting factors were the higher cost of Appforge, the requirement for a large runtime that consumed more hand-held memory, and the possibility that the application would be slower than a comparable C++ based application. Selection of the Palm Zire 21 hand-held device was based on its attractive price, size, weight, battery life, screen size and clarity, memory capacity and audibility of alarms. The Palm Zire 21 hand-held was compared to the Palm m130 and Palm Tungsten E which both offered more features such as color screens, slightly more compact dimensions, and minor weight savings. The most important decision factor was the remarkable difference in cost with the Zire providing superior value by virtue of it being half as expensive. During the pilot, users were queried as to the comfort, convenience and interaction with device. Neither weight, size, nor a black-and-white screen were mentioned as significant user concerns.

### 4.2. Software design

To enhance data entry efficiency, the software allowed linking of a response choice to a specific subsequent question. The participant thus followed a path of a branching logic tree based on prior responses. For example, a user was presented with an initial question, "My current activity can be described as: personal, transit between activities, direct patient care (hands-on, etc.)," which was followed by a question asking for more detail based on the response: "This *personal* activity involves: break/rest, eating, other (not work related)."

### 4.3. Data types

#### 4.3.1. Initial sign-on

Data, obtained once at the start of each participant's week-long study period, included participant job type (e.g., nurse or physician) and demographic information such as age, experience, and education. The design goals for participant response time are shown in Table 1.

#### 4.3.2. Daily sign-on

Data were obtained from each participant at the beginning of each work day (up to 7 times per study). Data included baseline information relevant to that day's activities (e.g., number of assigned patients) as well as subject activities since the previous work shift (e.g., quality of sleep in the preceding 24 h).

#### 4.3.3. Daily repeating surveys

These data were obtained multiple times each study day by prompting the participant randomly within consecutive 90-min intervals. Data included the specific current work activities, equip-

**Table 1**
Response rate results.

| | Response time design goal (min) | Number of responses obtained | Percent complete responses (%) | % Partial (incomplete) responses | Response time mean ± SD (median) in minutes |
|---|---|---|---|---|---|
| Weekly surveys | <10 | 638 | 100 | 0 | 6.60 ± 8.58 (4.13) |
| Daily sign-on surveys | <3 | 2305 | 100 | 0 | 2.27 ± 2.92 (1.78) |
| Repeating daily (random) surveys | <2 | 8435 | 73.3 | 11.6 | 1.48 ± 0.77 (1.38) |

ment in use, their emotional and cognitive state (e.g., mood, stress), and workload ratings.

Data were stored in a single "Responses" table on each hand-held. Question responses were stored as individual records, each with a unique question identifier and associated participant, day, and time stamp. Hand-held data, synced to the PC database, were then separated into different tables and columns with custom SQL data queries and table creation statements. This data structure allowed for question changes or additions without additional modifications or affecting database integrity.

### 4.4. Deployment

Each study site had a coordinator responsible for distribution, collection, and data downloads of the hand-helds. Our initial data download solution, a conduit to Hotsync devices over the web to a single centralized database, was thwarted by the hospitals' firewalls. Instead, we used built-in automated synchronization software (Hotsync for Palm OS) to transfer hand-held data to a dedicated desktop PC at each site. A custom Appforge conduit linked the hand-helds' data tables of questions and participant responses to the desktop PC's Microsoft Access database. If a modification to the question set was required, a single change made to the central "*Questions*" table was propagated to all devices hotsyncing to that database. Each week, study sites used a custom script to compress and email their database to the central database. Confidentiality was maintained by assigning each participant a unique anonymous numeric identifier, obtained via a separate centrally managed web-based participant database.

### 4.5. Iterative redesign

The tool design and development process occurred in a series of steps. Prior team experience with hand-helds to measure housestaff work [10] informed our initial sense of the structure of the survey tool related to work and of the maximum thresholds for response burden. We solicited input from a large interdisciplinary team of content experts to develop an updated tool that included new measures (NASA-TLX, mood, memory, etc.) and a new work survey applicable to nurses. The design team included psychologists, pharmacists, internists, hospitalists, nurses, health service researches, pediatricians, software developers, statisticians, human factors engineers. Iterative cycles of design and design team review yielded our first prototype. Then, we conducted ten cycles of iterative clinical testing of serial prototypes to assess response burden, usability and specific design features (e.g., battery life). The design team interacted on a weekly basis to review 4 clinician and site-leader input during the pilot tests.

To formally assess the feasibility of the Dynamic Hand-held Survey Tool and the user interface, a pilot study (beta-test) was performed. Physician subjects (first-, second-, or third-year medical residents) on the VA inpatient wards carried the hand-held tool for up to 6 consecutive days. Consenting housestaff were randomly sampled within consecutive 90-min intervals. An initial sign-on dialog captured demographic data and memory performance. Daily sign-on questionnaires queried subjects regarding clinical workload, sleep, and other factors. Eleven physicians were surveyed over 79 days, yielding 408 complete responses. Median time from hand-held prompt to initiation of response was 22 s, to complete an activity survey 16 s, and to complete the mood and task demand items 46 s. Thus, the overall response time was 84 s. The instrument successfully and efficiently captured vital descriptive data through daily sign-on dialogs (mean completion time 41 s); for instance, subjects reported a mean of 6.4 h of sleep and 7.6 patient census during the study period. Likert-scale ratings of mood and workload ranged across the full spectrum of the scales (0–9) for all items.

Based on the initial pilot studies, several design attributes and features were iteratively added or revised:

- To prevent users from inadvertently double clicking and thus skipping questions, we disabled the "Next" button until a selection was made;
- To reduce confusion, we added "High" and "Low" labels on slider input screens;
- To reduce response burden, confirmation notifications and requirements were limited to critical questions;
- To accommodate the small screen and prevent users from having to scroll, menu selections were limited to 37 characters and 9 total choices;
- To ensure application exit before device shut-off, a timeout function was added;
- Previously entered data fields (e.g., number of patients being covered) were pre-populated for the next survey to further reduce response burden;
- The hand-helds were stripped of all non-essential software;
- To prevent accidental program exit, the software locked all buttons, including "power-off", until the completion of a survey.
- To track infusion pump programming errors and link them to other study variables, a data entry screen was added. This allowed nurses to quickly indicate which specific pump (each pump in the clinical environment was tagged with a unique 5-digit identification code number) was being used at that time by that nurse. The screen was designed with four clickable columns of letters and a display text box which avoided the use of Palm Graffiti to enter pump codes. This was in response to user feedback that nurses were inputting invalid pump ID codes apparently due to simple clerical errors.

#### 4.5.1. Limited battery life

The hand-held's limited battery life became a major impediment to successful deployment. We chose to avoid solutions that required recharging the unit, because users' failure to comply would result in complete data loss. Battery life was most affected by the power required for display illumination during alarms, and while awaiting responses. Participants occasionally failed to indicate that they were going "off shift" and then left the device

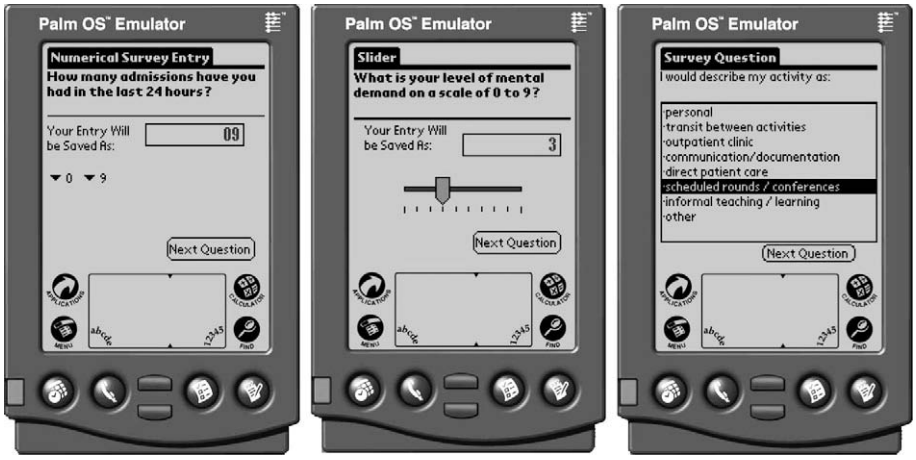


**Fig. 1.** Sample data screens showing different user interface interaction styles generated with Palm OS Emulator, or POSE, a Palm emulator made by PalmSource (PalmSource Inc., Sunnyvale, CA, http://www.palmos.com/dev/tools/emulator/).

on overnight. Therefore, users were prompted three times within each 90-min interval before the alarm was set for the next interval. If a user missed three consecutive notifications, the device reset to the start of the next workday, as specified by that user during their most recent daily sign-on. In addition, software was added to monitor battery life and display a warning at a specified threshold. When battery life reached this threshold, surveys ceased and a user message requested that the device be returned to the site coordinator.

### 4.6. User interface

The user interface presented challenges, the greatest being avoiding the need for users to enter text manually. Through iterative user-centered design, we created a number of different template screens that allowed users to make appropriate selections with only one or two clicks. To prevent the user from inadvertently clicking through two questions, we added a 'Next Question" button (disabled as a default) on each screen. After a selection, the user's choice was highlighted, and the "next" button was enabled. Whenever appropriate, data entry selections were pre-populated with default or expected options (Fig. 1).

## 5. Implementation

Initial software development took 6 months. Iterative pilot testing and redesign took an additional 2 months. Data were then collected over a 12-month period during which 304 clinicians (119 nurses and 185 physicians) at 3 hospitals provided 11,381 survey responses.

The hand-held tool proved reliable in capturing a large, multidimensional dataset from clinicians during real-time work. A high proportion of repeating survey prompts were completed (Table 1) suggesting that the participant response burden was acceptable in the context of a busy work place. The median completion time for each survey type was within the original design objectives (Table 1). Fig. 2 shows the distribution of response times, by type of provider (physician or nurse) for the initial sign-in (at the start of each participant's study period, A); daily sign-on (at the start of each study day, B); and the repeating surveys (randomly completed multiple times per study day, C). These results were obtained with virtually no training required of participants, who found the tool intuitive and easy to use. The resulting data have been analyzed to delineate the relationship between clinicians' psychological state, their task activities and workload, and the occurrence of medication errors [15–18].

## 6. Discussion

We developed a hand-held computer EMA application to facilitate collection of real-time data from busy inpatient physicians and nurses. Our key design goals were largely achieved. The device was reliable, data were captured consistently and accurately, and the response burden was sufficiently low to attain very high survey completion rates.

One limitation of this evaluation is the lack of a formal usability evaluation after the iterative design phase. However, clinicians reported that the tool was easy to use. For example, one clinician stated; "Screens and surveys have been working smoothly. The biggest concern as I expressed this morning is battery life." Overall, nurses' study participation and responsiveness were higher than that of housestaff. There was a decrease in the proportion of eligible housestaff (but not nurses or attendings) who were willing to participate in the study after their first weeklong participation. The decreased housestaff re-enrollment rate may have been due to their high clinical workload, the already frequent interruptions inherent in their normal work [19], and potentially, the intrusive nature of the study. Site leaders reported that some users "refused to take the palm pilot into meetings or otherwise into situations where they are going to be unwelcome interruptions." A smaller pool of eligible housestaff at the study sites necessitated more frequent participation requests than was originally intended. In contrast, nurses rarely refused to participate repeatedly, possibly due to the nature of their work and to their larger overall numbers. Overall, the high study acceptance rate (especially of nurses), the high rate of responses to prompts and the high quality of the data obtained [15,17,18] are consistent with good usability of the UI.

Some early concerns proved unfounded while use experience pointed to design improvement opportunities. Although the AppForge application[1] was slower than other software alternatives, there was no noticeable lag between screen transitions. Post-download data processing proved to be clumsy and time-consuming. This may have been improved by storing data into separate tables for each survey type (e.g., initial, daily, and repeating). The addition of

[1] *Authors' note:* Since the completion of this project, AppForge is no longer commercially available. However, all of the results and lessons learned are believed to be generalizable to hand-held-based EMA applications using alternative software applications.

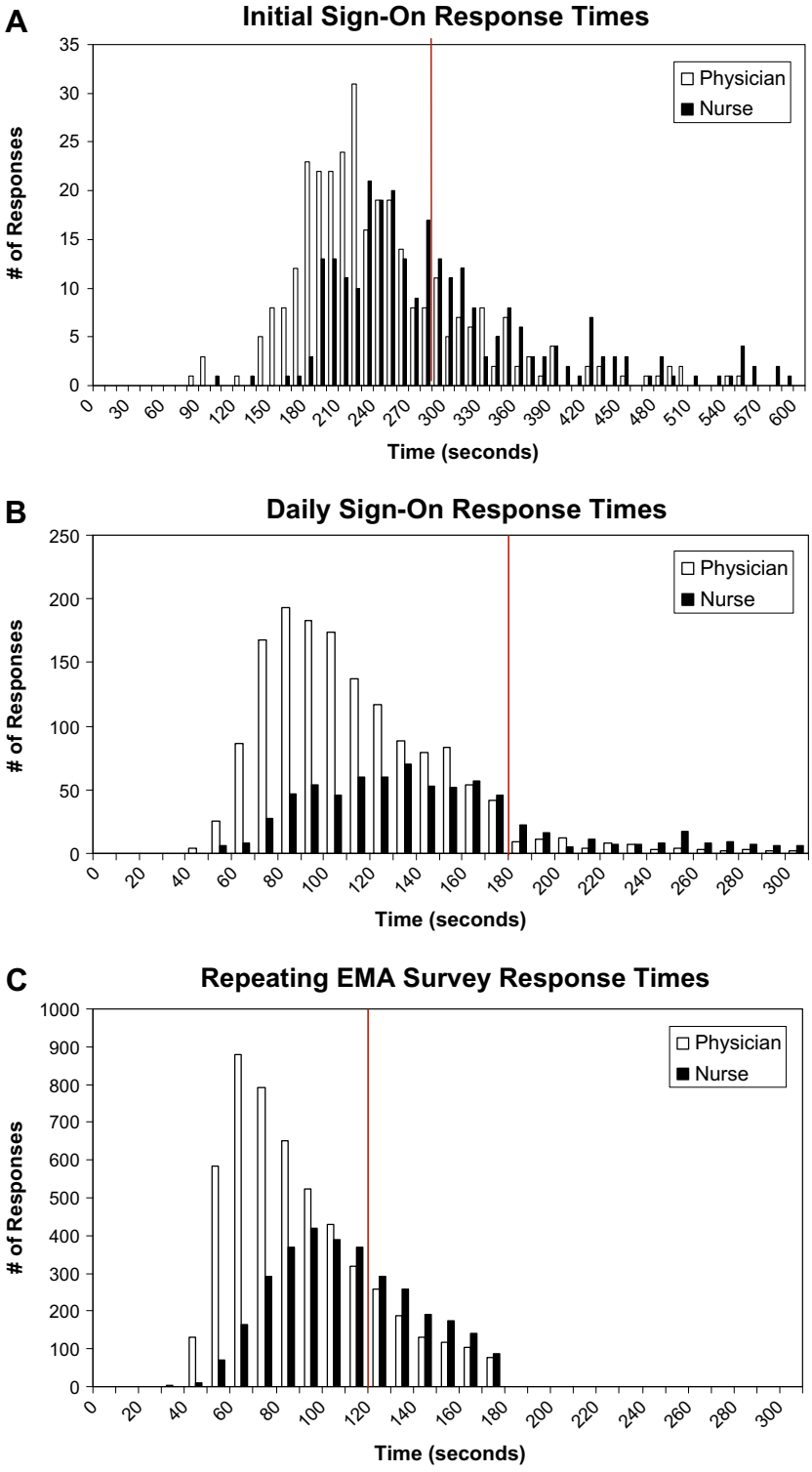


**Fig. 2.** The three histograms show the actual distribution of clinician participant (physicians and nurses) survey response times to the initial sign-on (A; design goal $\leqslant$300 s), daily sign-on (B; design goal $\leqslant$180 s) and repeating surveys (C; design goal $\leqslant$120 s), respectively.

color coding of selected headers and prompts could improve response time and interaction experience but would not have been supported by black-and-white screen hand-held computers.

This paper describes a user-centered design process and successful implementation of a hand-held instrument to collect data about clinicians' work behaviors, workload, stress, and psychological state. The results show that such an EMA tool can be effectively deployed to study clinician behavior in the actual work setting with low response burden and good participant compliance. A similar tool could have widespread application in the evaluation of informatics needs and adoption.

## Acknowledgments

This study was supported primarily by Grant UC1HS14283 (to T.R.D.) from the Agency for Healthcare Research and Quality

(Rockville, MD). Dr. Weinger's time on this project was also supported by Grants from the Agency for Healthcare Research and Quality and from Dept. of Veterans Affairs Health Services Research and Development. The technical assistance of Todd Kuykendall was greatly appreciated. Karen Ruberg and Kelly Vo provided invaluable administrative assistance. This study would not have been possible without the support and forbearance of the clinicians and administrators at our three study hospitals.

## References

[1] Kamarck TW, Debski TT, Manuck SB. Enhancing the laboratory-to-life generalizability of cardiovascular reactivity using multiple occasions of measurement. Psychophysiology 2000;37(4):533–42.
[2] Kamarck TW, Janicki DL, Shiffman S, Polk DE, Muldoon MF, Liebenauer LL, et al. Psychosocial demands and ambulatory blood pressure: a field assessment approach. Physiol Behav 2002;77(4–5):699–704.
[3] Collins RL, Morsheimer ET, Shiffman S, Paty JA, Gnys M, Papandonatos GD. Ecological momentary assessment in a behavioral drinking moderation training program. Exp Clin Psychopharmacol 1998;6(3):306–15.
[4] Shiffman S, Gwaltney CJ, Balabanis MH, Liu KS, Paty JA, Kassel JD, et al. Immediate antecedents of cigarette smoking: an analysis from ecological momentary assessment. J Abnorm Psychol 2002;111(4):531–45.
[5] Wegner KE, Smyth JM, Crosby RD, Wittrock D, Wonderlich SA, Mitchell JE. An evaluation of the relationship between mood and binge eating in the natural environment using ecological momentary assessment. Int J Eat Disord 2002;32(3):352–61.
[6] Smyth J, Wonderlich S, Crosby R, Miltenberger R, Mitchell J, Rorty M. The use of ecological momentary assessment approaches in eating disorder research. Int J Eat Disord 2001;30(1):83–95.
[7] Moskowitz DS, Young SN. Ecological momentary assessment: what it is and why it is a method of the future in clinical psychopharmacology. J Psychiatry Neurosci 2006;31(1):13–20.
[8] Freedman MJ, Lester KM, McNamara C, Milby JB, Schumacher JE. Cell phones for ecological momentary assessment with cocaine-addicted homeless patients in treatment. J Subst Abuse Treat 2006;30(2):105–11.
[9] Curran SL, Beacham AO, Andrykowski MA. Ecological momentary assessment of fatigue following breast cancer treatment. J Behav Med 2004;27(5):425–44.
[10] Dresselhaus TR, Luck J, Wright BC, Spragg RG, Lee ML, Bozzette SA. Analyzing the time and value of housestaff inpatient work. J Gen Intern Med 1998;13(8):534–40.
[11] AAMI Human Factors Engineering Committee: the human factors design process for medical devices (ANSI/AAMI HE-74:2001). Arlington, VA: Association for the Advancement of Medical Instrumentation.
[12] Nielsen J. Usability engineering. Boston: Academic Press; 1993.
[13] Wiklund ME. Medical device and equipment design: usability engineering and ergonomics. Buffalo Grove, IL: Interpharm Press; 1995.
[14] AAMI Human Factors Engineering Committee: human factors engineering guidelines and preferred practices for medical device design (ANSI/AAMI HE-75:2010). Arlington, VA: Association for the Advancement of Medical Instrumentation.
[15] Dollarhide AW, Rutledge T, Weinger MB, Dresselhaus TR. Use of a handheld computer application for voluntary medication event reporting by inpatient nurses and physicians. J Gen Intern Med 2008;23(4):418–22.
[16] Rayo MS, Smith Phil, Weinger Matthew B, Slagle Jason, Dresselhaus Timothy. Assessing medication safety technology in the intensive care unit. In: Proceedings of the human factors and ergonomics society 51st annual meeting; 2007. p. 692–6(5).
[17] Stucky ER, Dresselhaus TR, Dollarhide A, Shively M, Maynard G, Jain S, et al. Intern to attending: assessing stress among physicians. Acad Med 2009;84(2):251–7.
[18] Rutledge T, Stucky E, Dollarhide A, Shively M, Jain S, Wolfson T, et al. A real-time assessment of work stress in physicians and nurses. Health Psychol 2009;28(2):194–200.
[19] Gabow PA, Karkhanis A, Knight A, Dixon P, Eisert S, Albert RK. Observations of residents' work activities for 24 consecutive hours: implications for workflow redesign. Acad Med 2006;81(8):766–75.