Developing Design Guidelines for a Visual Vocabulary of Electronic Medical Information to Improve Health Literacy

XIAOJUAN MA*

Department of Computer Science and Engineering, Hong Kong University of Science and Technology, Clear Water Bay, Kowloon, Hong Kong

*Corresponding author: mxj@cse.ust.hk

Poor health literacy, or the ability to interpret and make judgments about one’s health, affects the effectiveness of healthcare services and people’s quality of life. We explored design principles on how to design visually salient, legible and interpretable medical graphics to promote comprehension and recall of electronic medical diagnostic information, an essential component of personal health records, for populations with limited health literacy. We first conducted a two-stage investigation to confirm that people can comprehend pictorial representations of medical conditions and treatments and improve their health literacy. We then designed a set of medical diagnostics graphics, based on the principles of information design derived from a large set of information graphics in various domains. Evaluated in a lab study, our medical diagnostic graphics were found to be capable of improving people’s perceived comprehension and recall, and were considered easy to understand and very useful for individuals with limited health literacy.

RESEARCH HIGHLIGHTS

- Medical graphics in electronic personal health records can improve people’s comprehension and recall of diagnostic information, and enhance their health literacy.
- People are good at recognizing simple symptoms and actions from medical graphics, but have difficulty in integrating pieces of information to make differentiation and judgment.
- Photo-realistic illustrations provide the most details to support interpretation, but can be disturbing; animations better depict procedure and change of status, but require patience in viewing; while clipart type of representations make good use of visual cues to convey details and process, and are quick to perceive.
- It is important to properly design visual features that highlight important facts, reveal details, demonstrate cause–effects, elicit terminology and measurements, e.g. color, line, symbols, text and numbers, frames, view and perspective, change of scale, style, as well as the presence and design of human character.

Keywords: personal health records; user centered design; user studies; crowdsourcing

Received 11 December 2014; Revised 7 May 2015; Accepted 24 June 2015

1. INTRODUCTION

Healthcare is closely related to well-being. Complete and accurate personal health information could be lifesaving, in situations ranging from medical complications that are treated by clinicians in different domains and different hospitals, to coordination in post-discharge care, to interactions with insurance organizations. A personal health record (PHR) is a tool that uses technology to access, track and share up-to-date information related one’s health, maintained by both patients and caregivers. PHRs provide centralized management of complex personal health information from many sources, and allow individuals to play a more active role in their personal healthcare, enabling higher self-control, independency and better decision support (Tang et al., 2006).
From an organizational perspective, a PHR could reduce the cost of information transfer, and improve the quality of service when integrated with an electronic health record (EHR) system maintained by each healthcare provider. Hence, there has been considerable global interest in promoting the adoption of PHRs (Paglari et al., 2007). However, despite the strong consensus on the perceived benefits of a PHR, and a notable upsurge in campaigns by government and private sectors to encourage the deployment of PHRs, a sea change has yet to take place (Tang et al., 2006). After Microsoft reported that efforts on to make its PHR application HealthVault a profitable venture were abandoned (Jack et al., 2009), Google announced the closure of its Google Health service at the end of 2011 (Google Health, 2011).

Several obstacles to the adoption of PHRs have been identified. PHRs rely heavily on individuals’ ability to obtain, process and understand health information and services (Ratzan and Parker, 2000), as well as their motivation for manually recording data (Lober et al., 2006; Tang et al., 2006). This is particularly true for records, including medical conditions, laboratory results, diagnoses, medication, home care plans, etc. If a patient cannot make sense of information presented by the physician, it is unlikely that the information will subsequently be recalled and entered into the PHR. In this paper, we focus on strategies to improve the ease-of-use of PHR by improving patients’ comprehension and recall of diagnosis-related information over the potential impedance of insufficient health literacy.

Health literacy is defined as ‘the degree to which individuals have the capacity to obtain, process and understand basic health information and services needed to make appropriate health decisions’ (National Institute of Health, 2014). Low health literacy is associated with inadequate experience and education about health knowledge, and can be compounded by factors like low literacy, cultural differences and insufficient language proficiency (Andrus and Roth, 2002). Surveys showed that up to 80% of the patients in public hospitals have insufficient health literacy (Williams et al., 2002). Particularly, lower health literacy is associated with all racial and ethnic groups, people with less education, and the elderly (U.S. Department of Health and Human Services, 2015). Poor health literacy greatly impedes patient–physician communication, patients’ health practices at home and their health outcomes (Williams et al., 2002). PHRs could overcome health literacy barriers and improve the quality of patient–physical communication.

Many strategies have been implemented to accommodate for inadequate health literacy, including interpreters, cultural competency training and coordination with traditional healers, community health workers, and family members in patient care (Riddick, 1998). These strategies are costly, and any involvement of a third-party translator raises privacy concerns. Another method has been introduced: adding pictorial illustrations to medical information (Andrus and Roth, 2002; Williams et al., 2002). The role of pictures in improving communication in medical care has been thoroughly investigated in previous research. Pictures have been proven to have positive impact on people’s attention to health materials, comprehension and recall of health information, and the likelihood of adherence to medical instructions (Houts et al., 2006). The majority of illustrated health information is available only in printed forms. In addition, different visual vocabularies have been used in the graphical health information, e.g. Houts et al. (2006). There is a lack of standard design of medical graphics in this domain, and a comprehensive understanding of how different visual features may affect people’s perception and interpretation.

Therefore, we propose the use of information and communication technologies, e.g. adaptive multimedia display, to expand the benefits of a standardized graphical medicine vocabulary to clinical diagnostic information used in PHRs. A standardized system can facilitate the exchange of a wide variety of health information between clinician and patient, ranging from screening questions to treatment instructions. The system can help to regulate the diagnostic process and reduce errors introduced by miscommunication; furthermore, it could serve as an interface to create new entries for updating the patient’s PHR. In the future, such design can be applied to other healthcare settings which may suffer from health literacy problems, such as triage, pediatrics, speech therapy, medical aids in remote areas, self-diagnoses website, etc.

In this paper, we present our design, development and evaluation of a visual vocabulary of medical information employed in a standardized digital graphical diagnostic template, which is an essential component of PHRs. We first investigated health literacy problems in people’s daily medical care practices via a survey which consisted of a questionnaire and an ailment–treating matching test (Section 3). Results suggested potential areas that may benefit from the use of medical graphics. We then conducted a crowd-sourced labeling study to explore how well people could identify medical information presented in different visual forms, e.g. cliparts, illustrations, line drawings and animations (Section 4). Next, in Section 5, we gathered a set of commonly used graphical information presentations in different domains, and extracted salient features. We constructed taxonomy of notifications that can reliably convey critical elements in diagnostic situations. Later, we developed a visual vocabulary for illustrating health information related to four diagnostic scenarios selected by physicians. We conducted a controlled study of documenting diagnostic communication in a PHR with and without the medical graphics we created, simulating the actual context of use (Section 6). Results verified our predictions that the standardized digital graphical diagnostic template could improve individuals’ perceived comprehension and recall of health information discussed. From these findings, we generated a set of design recommendations for creating standards for graphical diagnosis information.
2. BACKGROUND WORK

Our related work spans research in health literacy, PHRs and EHRs, icon design as well as graphical medical information.

2.1. Health literacy

The health literacy problem is prevalent. It is estimated that nearly half of the population in the USA encounter difficulties in comprehending and utilizing health information (National Institute of Health, 2014). People with the greatest medical care needs, including elders, people with low income, minorities, immigrants and people with chronic diseases (Glassman, 2011) are often lower in their health literacy level. As a result, a patient may experience physician–patient communication breakdown during the diagnosis process, may fail to follow medication dosage and schedule instructions, may miss follow-up doctor’s appointments, and be less likely to keep track of information using a PHR.

Health literacy is more than reading comprehension, requiring analytical and decision-making skill and their use in health situations (National Institute of Health, 2014). Many standards for testing health literacy have been developed (Baker et al., 1995; Davis et al., 1993; Weiss et al., 2005; Wide Range Inc., 1993). These tests aim to evaluate different aspects of health literacy, for instance, how well people can hold discussions about their medical conditions, search and process health information included in written and/or visual materials, perform calculations and use tools for family or personal care.

In this paper, we focus on people’s ability to comprehend and recall the most frequently updated information in a PHR—clinical diagnostic communication. This information may come from multiple health providers, and is critical in insurance billing purposes. If there were a way to assist diagnostic communication processing across the health literacy barrier, it would be easier for patients to adopt and profit from PHR.

2.2. Strategies to overcome health literacy problems in the promotion of PHR

Given the benefits of PHRs for patients, medical care providers and insurance providers, many have encouraged the deployment of PHRs. Prior work suggested that education on health management techniques should start in primary school and continue throughout different levels of education, so that potential users have better knowledge and acceptance of how to manage personal health using technologies (Tang et al., 2006).

For patients who have missed early opportunities to get educated on handling PHR with technologies, the quality of information exchange during the diagnostic process is critical to their participation in personal health management. To this end, healthcare providers have made attempts on training physicians on information communication skills (Stewart, 1995) and providing patients with sufficient educational materials (Kaplan et al., 1989).

While important, the effects of these attempts may vary from person to person due to factors such as literacy and experiences. Therefore, medical graphics along with means to interpret them have been employed, as a low-cost method without privacy concerns to make medical information more comprehensible for those with low health literacy (Kaplan et al., 1989).

2.3. Visual communication

People can potentially benefit from the use of medical graphics because visual communication is universal. We witness it every day through signage, maps, diagrams, airline safety cards, assembly manuals and other information graphics, and thus we are more or less trained to recognize visual information in these styles. Existing literature indicated that several design features are especially effective in conveying medical information. For example, contours in line drawings can depict the most essential details of sizes, shapes and locations of their subjects (Rubin, 1958). Clipart or cartoon like representations can further utilize visual cues such as colors, textures, shading, symbols and letters to facilitate graphical communication strategies—realism, simplification, exaggeration and symbolism (McCloud, 2006). In addition, visual stimuli with multiple frames (e.g. animations) have been found, if carefully designed following the Congruence Principle and Apprehension Principle, to better represent sequencing, timing and changes over time (Tversky and Morrison, 2002).

2.4. Graphical medical information

Graphical information that adopts the useful visual cues listed in the previous subsection has proven to have a positive effect on promoting communication in healthcare (Houts et al., 2006; Katz et al., 2006). In particular, graphics can direct attention to medical information. It has been shown that people are more likely to read and follow handouts with a cartoon embedded in the text (Berkman et al., 2004). Graphics can improve understanding. Prior work showed that pictures (Morrow et al., 1998; Riddick, 1998), animations and videos (Lee, 2003) could help patients comprehend complex medical terminology. Furthermore, graphics can provide cues for recalling information such as doctor’s instructions (Lee, 2003; Morrow et al., 1998). Previous work has also raised other important issues regarding the use of medical graphics. For instance, the style of graphics, e.g. photos versus colored drawings versus black and white line drawings (Berkman et al., 2004), and the content of the graphics, e.g. pictures of an entire person versus pictures of a body part (Morrow et al., 1998) may affect viewers’ comprehension.
Medical graphics have also been shown to improve physician–patient communication in a variety of cultures, particularly in identifying symptoms (Lim, 2010; Medhi et al., 2007). In the domain of PHRs, researchers have studied the use of graphics to enhance comprehension of medication instructions (Dowse and Ehlers, 2005; Katz et al., 2006; Riddick, 1998), discharge instructions (Austin et al., 1995; Delp and Jones, 1996) and education materials (Leiner et al., 2004; Michielutte et al., 1992; Moll, 1986). Medical graphics have yet to be investigated in other medical contexts such as diagnosis, laboratory tests and results, diagnosis and treatment. These data are even more critical to the reliability and usefulness of a PHR than a list of possible symptoms.

Additionally, many digital medical graphics have not yet been broadly disseminated to the public (Dowse and Ehlers, 2001). There is a lack of investigation about what design features are the most salient and effective to illustrate health information, given the diversity in people’s visual literacy (Hämeen-Anttila et al., 2004). As Norman (1988) suggested, it is critical to ensure that the user’s model developed from the system image—e.g. interface and graphics used—corresponds to the design model. Therefore, our work fills this gap by developing design recommendations for standard graphical elements to communicate health-related information.

In this paper, we aim to explore the problem space of diagnostic communication barriers in healthcare, and to provide the theoretical basis for the design of a pictorial language in medical care across communication barriers, and to promote the quality of healthcare services while reducing the cost. The research questions that we would like to address are:

Q1: How do people with different health literacy practice healthcare in daily life, especially with the support of medical graphics? Understanding the user’s model (Section 3)

Q2: How can an effective visual vocabulary for PHR support system be designed, given the complex process of diagnostic examination? Materializing the design model in system images (Sections 4 and 5)

Q3: Can pictorial support developed based on our guidelines improve doctor–patient diagnostic communication (i.e. comprehension, and recall)? Investigating whether system images are consistent with the design model and operate according to the user’s model (Section 6).

Figure 1 provides an overview of how the research activities presented in different sections help address these three research questions.

The outcome of this research includes design guidelines for medical graphics that can be generalized to all kinds of information in PHRs, as well as an online graphical diagnostic communication system that we are in the process of developing. We aim to improve patients’ experiences without loading extra burden onto physicians, which is one of the key concerns of healthcare providers about the use of PHRs.

3. INVESTIGATION OF PRACTICES IN MEDICAL CARE AND THE USE OF MEDICAL GRAPHICS

To understand how health literacy may influence people’s knowledge, perception and behaviors, and what the potential areas are where medical graphics can benefit users, we carried out a preliminary investigation. We first gathered information on personal practice and experience with medical care via a questionnaire with different populations. Then, we conducted an ailment–treatment matching test to explore how people made use of medical graphics when available.

3.1. Questionnaire on daily healthcare practices

To actually understand how health literacy may influence people comprehension of health information (where they occur; in what format; for what reason etc.), and what are the potential areas where medical graphics could benefit the...
readers, we carried out a preliminary survey with 20 participants which included a questionnaire and a within-subject ailment–treatment matching study.

3.1.1. Participants
We recruited 10 native English speakers and 10 native Chinese speakers. Within the native speakers’ group, half were female, all between the age of 20 and 65. Four non-native speakers were females. Two people were in the 31–40 age group, and the rest were between 21 and 30. All participants had at least college degree. The non-native speakers were graduate students or visiting scholars from China who had never lived in an English-speaking country prior to their visit to the States. We chose native Chinese speakers as the representative group with limited health literacy because of their relatively low English proficiency and different medical practices in the home country. They also form one of the biggest non-native English-speaking communities in the local area, where several universities together introduce a large body of international students and scholars. According to the physicians and staff of the local hospital that we interviewed, it is challenging to serve a population with big culture and language disparities. All participants completed a self-assessment of their knowledge of medical care (scale 1–5 per question). All the non-native speakers had a score lower or equal to 3 for every question, although their oral English proficiency varied from Level 2 ‘very simple conversation’ to Level 4 ‘with a fair degree of fluency’ (O’Loughlin, 2001). In contrast, all native speakers were above 3.

3.1.2. Questionnaire results
We asked participants to fill out a paper-based questionnaire1 regarding their knowledge, attitudes and experience with medical care, in addition to demographics information and self-assessment of English proficiency:

(i) How much medical terminology do people know (1/no idea at all ∼5/very knowledgeable)? The average score for native speakers was 3.5 (standard deviation SD = 0.85), and non-native speakers was 2.25 (SD = 0.79).
(ii) How well can people describe their symptoms (1/very poorly ∼5/very well)? The average score for native speakers was 3.8 (SD = 0.63), and non-native speakers was 2.5 (SD = 0.97).
(iii) What is their approach when feeling sick? Eight out of 10 non-native speakers chose to go to the doctor if they felt sick for more than 2 days. On the contrary, only two native speakers chose to see a doctor, while five took home remedies, and three asked family members for help.
(iv) What medications do people keep at home? Nineteen participants keep medications at home. About 83% of the Chinese participants had mainly traditional Chinese medications. The most common commercial medicines kept were Tylenol (7), pain reliever (6, e.g. Advil), Aspirin (6), allergy medicines (4), First Aid Kits (4).
(v) How do they find medications? Eight native speakers tried to figure out what to take on their own, and five of them got prescriptions from doctors. On the contrary, seven non-native speakers relied on a doctor or family member knowledgeable in medical care. Only three non-native speakers said they would try to find medicines on their own.

3.2. Ailment–treatment matching test with and without medical graphics
The ailment–treatment matching test was carried out with the 20 participants recruited after they completed the questionnaire. The goal was to reveal the health literacy and cultural competency problem related to medical information comprehension, and evaluate the efficacy of medical graphics support.

3.2.1. Matching test design and interface
The basic task of the ailment–treatment matching test was to assign 18 medical treatments to 20 most common illnesses (Table 1) that they are most likely to be used for on a web interface. Note that a treatment may be suitable for multiple illnesses, and an illness may have multiple treatments. We explained the task to the participants, and instructed them to complete an example to get familiar with the interface. Participants could revise the answer before submitting it. We compiled the list of illnesses and treatments based on interviews with doctors in a local university medical center and the statistics pulled from the previous year’s record. In the first round of the matching test (Fig. 2a), we provided text descriptions for both the ailments and the treatments which came from online resources such as the NIH MedlinePlus website (MedlinePlus, 2013). The descriptions were (i) tailored to reflect only those facts that were related to the matching task; (ii) modified to embed hints such as ‘used as the major treatment for the following symptom’ to reduce ambiguity; (iii) changed to provide both professional medical terms and common ‘nick names’ to simulate daily practices; and (iv) adapted so that different words were used to describe symptoms in the treatment and ailment to prevent solving the matching task based on matching up similar text. Note that such task design limited the scope of choice and reduced information complexity. In addition, we deliberately removed some shortcuts at the syntactic level to force people to process information at the semantic level. Therefore, the task was different from what people may encounter in a natural setting.

In the second round (Fig. 2b), in addition, we collected 185 icons (40 line drawings, 107 clipart files and 38 illustrations) and 20 animations (specifically for the ailments).
Table 1. Left is the list of 18 treatments and 20 most important diagnoses of ailments numerically from a local university medical center; note that no treatment was for common cold.

<table>
<thead>
<tr>
<th>Treatment</th>
<th>Targeted ailment</th>
<th>Native English speakers</th>
<th>Non-native English speakers</th>
</tr>
</thead>
<tbody>
<tr>
<td>Amoxicil</td>
<td>Lyme disease</td>
<td>Clearasil</td>
<td></td>
</tr>
<tr>
<td>Calamine lotion</td>
<td>Poison ivy</td>
<td>Dramamine</td>
<td></td>
</tr>
<tr>
<td>Clearasil</td>
<td>Acne</td>
<td>EpiPen</td>
<td></td>
</tr>
<tr>
<td>Dramamine</td>
<td>Motion sickness</td>
<td>Ex-lax</td>
<td></td>
</tr>
<tr>
<td>EpiPen</td>
<td>Peanut allergy</td>
<td>Gatorade</td>
<td></td>
</tr>
<tr>
<td>Ex-lax</td>
<td>Constipation</td>
<td>Insulin</td>
<td></td>
</tr>
<tr>
<td>Gatorade</td>
<td>Food poisoning</td>
<td>ORIF</td>
<td></td>
</tr>
<tr>
<td>Gatorade</td>
<td>Gastroenteritis</td>
<td>ProAir HFA</td>
<td></td>
</tr>
<tr>
<td>Calamine lotion</td>
<td>Heatstroke</td>
<td>Prozac</td>
<td></td>
</tr>
<tr>
<td>Open reduction and internal fixation</td>
<td>Ankle fracture</td>
<td>ORIF</td>
<td></td>
</tr>
<tr>
<td>Pepto-Bismol</td>
<td>Dyspepsia</td>
<td>RICE protocol</td>
<td></td>
</tr>
<tr>
<td>ProAir HFA</td>
<td>Asthma</td>
<td>RICE protocol</td>
<td></td>
</tr>
<tr>
<td>Prozac</td>
<td>Anxiety, depression</td>
<td>Statin</td>
<td></td>
</tr>
<tr>
<td>RICE protocol</td>
<td>Ankle strain</td>
<td>Tecnu</td>
<td></td>
</tr>
<tr>
<td>Relenza</td>
<td>Influenza</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Statin</td>
<td>Cardiac disease</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Tecnu</td>
<td>Poison ivy</td>
<td></td>
<td></td>
</tr>
<tr>
<td>TUMS</td>
<td>Dyspepsia</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Zyrtec</td>
<td>Hayfever</td>
<td></td>
<td></td>
</tr>
<tr>
<td>–</td>
<td>Common cold</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Right are lists of successfully assigned treatments by all participants in each group; the ones in bold were answered correctly when medical graphics were available.

Figure 2. Web interface of the ailment–treatment matching test: (a) given text description only in the first round; (b) given text description with medical graphics in the second round.

from the Internet to illustrate critical words and phrases in the descriptions. When participants move the cursor over the underlined terms, the associated graphics will pop up. We asked participants to clarify whether each of their assignment in the second round was affected by the medical graphics they had viewed.

3.2.2. Matching test results and analysis

Table 1 lists the successful matches by all people in each participant group in the first and second rounds of the ailment–treatment matching test, respectively. Overall, native English speakers performed significantly better than the non-native speakers in terms of the number of correct assignments.
Developing Design Guidelines

(a) Improved matching results based on text alone versus with medical graphics involved, as well as the number of clicks (with standard errors) to view pictures in the descriptions (b) by native and non-native English speakers.

Figure 3. Improved matching results based on text alone versus with medical graphics involved (a), as well as the number of clicks (with standard errors) to view pictures in the descriptions (b) by native and non-native English speakers.

(F(1,18) = 49.59, P < 0.05). It was because past experience affects a person’s knowledge about medical care. The three participants above age 40 knew more about illnesses and had used or heard of various treatments. Younger Americans, in contrast, mentioned that they had heard of some of the medications on commercials, but could not remember what they are for. Non-native English speakers who had little experience with Western medications only knew medicines from their home countries. Much of their knowledge of illnesses and especially western medications came from American television shows and movies.

Figure 3a shows the number of assignments that were correct in the second round but not in the first round, with indication of the basis of judgment by each group. The interaction effect between group and the availability of medical graphics was significant (F(2,36) = 10.38, P < 0.05). This happened because the native English speakers had pre-conceived notions about what treatments were the best despite contradictions with what the text presented. Non-native English speakers achieved significantly more correct matches with the help of pictures for than native speakers (F(1,18) = 7.67, P < 0.05). This implies that visual representations may be a good media to promote one’s health literacy, when words are not as effective on conveying the information. On average, non-native speakers clicked on pictures 50 times more in each study, indicating that they used more assistance from the pictures than native speakers (F(1,18) = 10.25, P < 0.05; Fig. 3b). According to participants’ self-assessment of English proficiency in the demographics questionnaire, those with relatively poor English skills were more likely to view medical graphics.

We also observed participants’ attitudes and behaviors toward the medical graphics. The animations were too long (~5 s each). Our participants did not have the patience to watch the entire animation and frequently quit after viewing only five frames. Many non-native English speakers used medical graphics as a quick check on their understanding of an illness or treatment, especially, those who could communicate general information in English—Level 3 and 4 in (O’Loughlin, 2001). Assigning treatment to one or more out of 20 illnesses can be heavy on cognitive load, as Miller’s Law (Miller, 1956) suggests that generally humans can hold 7 ± 2 objects in the working memory. Medical graphics provided information on the type of disease (skin, digestion or heart problem) at a glance, bypassing the need to read the text descriptions in details. For example, three participants who initially mistook ‘heatstroke’ for ‘heart-stroke’ quickly realized that it was not a cardiac disease when they saw the visual representation. Also, people looked for visually similar items in graphics when doing the matching. For example, many participants checked the existence of poison ivy leaf across all the illness graphics when they saw it in the calamine lotion treatment image.

3.3. Discussion on the questionnaire and the matching test

Our investigation on the common practices of medical care of people with different levels of health literacy and their possible use of medical graphics revealed some interesting implications:

(i) People with more medical care practices, mostly native English speakers, were likely to match treatment based on superstition/prior beliefs from life experience (e.g. what they used at home), which may not be correct. For instance, 8 out of 10 native speakers believed that Amoxil, an antibiotic, should be used to treat viral infections such as colds and flu.

(ii) People do not always read carefully. For example, we put down in the description for food poisoning that ‘The main treatment is putting fluids back in the body’, but 75% of the people still selected antibiotics.

(iii) Some of the native English speakers did not look at the iconic illustrations carefully in the second round. Four native speakers indicated in the post-study interview that even though they were not certain about the answers, given their language skills, pictures would not bring more benefits to their comprehension of the information. Some people who did look at the pictures...
felt that certain illustrations made the information even more confusing.

These findings reveal several goals that good graphical health information needs to achieve: (i) visual saliency, so that people who tend to ignore such help have a chance to be exposed to it; (ii) legibility, and (iii) interpretability. As a next step, we explored whether existing medical graphics have satisfied these requirements.

4. EXPLORATION OF DIFFERENT FORMS OF MEDICAL GRAPHICS

There have been a big variety of visual representations of health information created, and they are widely available over the Internet. We carried out a crowd-sourced medical graphics labeling study to assess the effectiveness of existing designs. We aimed to evaluate the ability of ordinary individuals to interpret medical graphics, and gain a better understanding of the design space.

4.1. Design of the crowd-sourced medical graphics labeling study

We conducted the study via Amazon’s Mechanical Turk (mTurk)² to obtain human labels for the medical iconic representations collected from the web. The goals were (i) verify if people can identify the medical concept based on the icons and animations; (ii) help to pick a good representation for the chosen medical terms and phrases; and (iii) suggest rules for selection and creation of medical icons and animations. We collected 340 graphics, including line drawings, clipart files, illustrations and animations from the Internet that illustrated ~200 medical terms and phrases.

We recruited 156 mTurk workers to participate in the study. Each graphical representation was labeled by 10–20 people. We restricted the participants to US citizens using mTurk’s filtering mechanism, and mTurk ensured that each participant could not repeatedly label the same image. According to the demographics information we collected, the participants were between the age of 20 and 65, and 58% were female. All had finished at least high school education. In each task, 10 medical graphics were displayed one after another. We instructed the participants to describe the symptom(s) depicted in the picture, identify the medical diagnosis and explain the possible cause(s) of the symptom(s)/ailment (Fig. 4). The interface recorded their answers, usually in short phrases or sentences. The use of mTurk allows quick, rather cheap access to human judgment, but the results can be noisy as we have no control over participants’ behavior. Therefore, we manually checked the results and excluded workers whose responses were of poor quality, that is, the response was meaningless (e.g. ABC) or

irrelevant to the given picture (e.g. hello). The final submission acceptance rate was 98%.

4.2. Labeling study results and analysis

We computed the accuracy rates (% of people getting the answer completely right—using the same expression or synonym) for each of the three questions of all medical graphics. We found no significant difference in terms of average accuracy per representation given different visual styles. About 87% of the medical graphics had an accuracy rate averaged over the three questions >0.5, meaning that they could be correctly recognised by half of the participants (Table 2). More specifically, regardless of the style, people were good at recognizing graphics of pain at different locations (seven representations all higher than 0.9), panic and fatigue, heart disease (e.g. high blood pressure, heart attack) and common flu/cold like symptoms (runny nose, fever, headache and sore throat). We found that people could make similar mistakes in animations and in static graphics. For example, in both representations, about a third of the people mistook ‘ankle sprain’ for ‘ankle fracture’.

We further analyzed the medical graphics associated with 25 most common illnesses according to the records in a local university medical center. We computed the percentages of people identifying the problem, describing the right symptom(s), and coming up with the correct medication term of the ailment, respectively (Table 3). In general, participants were better at identifying symptoms than causes or diagnoses. At least 60% of labelers correctly described the symptoms for 23 (except aphthous and Lyme disease) out of the 25 ailments, versus 18 ailments for cause and 15 for diagnoses. This was because an illness may come with multiple symptoms and the same symptom may occur in multiple ailments. For instance, sneezing and runny nose are common symptoms for cold, flu and hayfever, and thus people mistook these three diseases for one another if they failed to piece together various clues, e.g. the external and internal cause as well as potential treatment.

| Table 2. Percentage of medical graphics at each accuracy rate level (left), and the mean and standard deviation (SD) of accuracy given different styles of visual representations (right). |
|---|---|---|---|
| Accuracy | Percentage of medical graphs | Style | Mean accuracy | SD |
| 1.0 | 16% | Animation | 0.67 | 0.27 |
| 0.9–1.0 | RICE protocol | | | |
| 0.8–0.9 | 15% | Line drawing | 0.58 | 0.29 |
| 0.7–0.8 | 9% | | | |
| 0.6–0.7 | 12% | Clipart | 0.69 | 0.27 |
| 0.5–0.6 | 22% | | | |
| <0.5 | 10% | Illustration | 0.72 | 0.24 |
| 0.0 | 3% | | | |

4.3. Implications for the development of diagnostic graphics

Our preliminary investigation (Section 3) showed that graphics can benefit people with low health literacy in understanding symptoms and treatments related to the diagnostic process. We further analyzed the reasons why people made mistakes in the medical graphics labeling study by coding all the overlooked or confusing features associated with the errors. Given the vast design space, it is important to derive guidelines for the design of health information graphics, with particular emphasis on:

(i) Features that are used to highlight important facts which will attract people’s attention to critical objects and locations, and thus can improve comprehension and recall of information related to identification and location/direction. For example, 40% of the labelers mistook ankle sprain for ankle fracture because they did not note that the impairment occurs in the ligaments rather than the bones.

(ii) Features that are used to reveal detail which will make it easy for people to map mental states to real-life
situations, and thus can enhance understanding and disambiguation. For instance, although 80% of the labelers correctly identified the symptom of pain in joints based on lightning-like symbols, fewer than half of the people associated it with arthritis because many of the given graphics failed to illustrate the internal inflammation.

(iii) Features that demonstrate cause–effect information which will help to explain conditions and relations which are often embedded in diagnostic information, and thus can improve people’s comprehension and recall. For example, fewer than 30% of the participants successfully differentiate influenza from common cold, as most people could not tell that the symptoms were caused by seasonal flu virus. In contract, all labelers were able to recognize hayfever which also has very similar symptoms, give the depiction of pollen as a strong indicator.

(iv) Features that elicit identification and measurement which will allow people to make visual comparison, and thus promote memorization of information related to dimension, degree and scale. For example, the illustrations of high blood sugar levels helped 75% of the labelers picked out diabetes.

5. DEVELOPMENT OF A VISUAL VOCABULARY OF MEDICAL DIAGNOSTIC INFORMATION

To develop a set of legible, interpretable and visually salient features for depicting health information, we extracted the most frequently used design features from existing examples of information graphics for the four aspects identified in the guidelines above, and evaluated them with a professional designer and potential readers.

5.1. Collection of information graphics for feature extraction

We used three domains of information graphics to extract common design features: 24 airline safety cards (from 20 countries), 15 assembly manuals (for furniture and equipment from resources like IKEA, Target, Safco, etc.) and 10 medical kits/pictograms. We selected the first two domains because these pictorial illustrations were universally used for readers with different cultural and language background, and they cover most of the communication elements required for diagnostic information specified in Houts et al. (2006). The medical graphics datasets include standard ones such as the USP-Pictograms which have been used in prior studies (USP, 2011), emergency instructions from the United States ready.gov (Ready.gov, 2011), the pictogram library from The Risk/Benefit Assessment of Drugs-Analysis and Response (RAD-AR) Council of Japan (RAD-AR, 2011), A.D.A.M. medical images A.D.A.M. (2011), as well as graphics from communication cards, travel kits and other research/commercial applications. We fetched all these graphics online.

5.2. Feature selection

We marked the appearance of features, recorded details and intended functionality for each, and computed the frequencies of appearance. From the information we gathered, we developed a set of design features based on (Mijksenaar and Westendorp, 1999), which were related to the four aspects (i.e. highlights, details, cause-and-effect and identification and measurement) specified in the guidelines from Section 4. The main features included:

(i) Color. The color scheme of the main content is generally desaturated. Bright colors are used for highlighting; if the image is black and white, pure black or thick black lines are used for highlighting. Red is reserved for warnings and ‘don’ts’, while yellow indicates that caution is required. Important elements in graphics are often colored orange or yellow—high chroma, high key hues. Red arrows frequently appear in instructions on action and direction/orientation.

(ii) Line. Solid lines are mostly used as auxiliary lines or call-out lines; dashed lines serve either as an indicator of past/conditional situation or as auxiliary lines; and heavy lines are used to highlight the outline of an item, location or auxiliary symbols.

(iii) Symbols. Arrows indicate direction, scope or cursors; a cross is a universal ‘no’ symbol (sometimes \ or ⊘), and occasionally used a marker for location (‘∗’ is another variation); basic shapes like a circle or rectangle are used to mark a particular item or location; a cursor could be a hand or an arrow symbol to point out certain item or location; and ‘!’ sometimes appears in warnings.

(iv) Text and numbers. Text is used for specific names or detail on warnings, sequences or actions. Numbers are used to identify locations or items, to indicate measurements or as indices for sequences and movement.

(v) Multiple frames. Multiple images may be combined to show steps in sequences and movements, different situations or outcomes (e.g. in warnings), and cause and effect relationships.

(vi) Views and Perspective. An exploded view is mainly used to show composition and connection; a front view is adopted mostly for identification, and sometimes for action and location; a side view is mostly used for location, and occasionally for identification in combination with a top view; a diagonal view is the most common, appearing in identification, composition, action, location, movement and connection.
Developing Design Guidelines

5.3. Design of diagnostic medical graphics

It is impossible to evaluate the performance of design features in particular. Designs, instead, communicate holistically through a combination of all of their features. Therefore, in order to assess the effectiveness of the design guidelines, we created a set of medical graphics which grew from the feature analysis of the set of information graphics we had gathered. We undertook an iterative, user-centered design process. A first round of designs was sketched by hand by a professional designer and evaluated by three additional designers and several test readers. From there, we standardized a visual language and color palette, and refine the set of medical graphics to the final list. We created the final static designs using Adobe Illustrator. Figure 5 shows examples of our medical graphics, depicting a symptom, a diagnosis and a treatment respectively.

6. ASSESSMENT OF DESIGN FEATURES IN MEDICAL GRAPHICS IN A DIAGNOSTIC SETTING

Owing to the restrictions of carrying out studies with patients in clinical settings such as concerns about privacy and

In summary, color, line and symbols are features for highlighting important facts. Change of scale and (partial) photo-realistic style are used to reveal details. Multi-frames with or without arrows and number indices can demonstrate process as well as cause–effect information. Text and numbers embedded in the graphics often indicate specific terminology and measurement.

Figure 5. Example of medical graphics we developed, left to right: dizziness when sitting—symptom, influenza—diagnosis and cough hygiene—treatment.

(vii) Change of scale. Zoomed views like fish eyes and magnifying glasses are often used to call-out details of cause and effect, composition, action, location/orientation and connection.

(viii) Style. Cartoon style is the most frequently seen style, because it is simple and casual. A clean information graphic style is normally used in warnings, composition and connection where the details are important. Photo realism is rarely used in this domain. It is reserved for customization or depicting a scene in real life.

(ix) Presence of People. Parts of the body are shown that are related to relevant situations and actions. Indication of emotion is rare in assembly manuals, and relatively uncommon in airline safety cards (mostly is a smile in safety instructions or some changes of facial expressions as a movement goes on). However, medical instructions often use emotions to indicate health conditions (e.g. a pain scale).

(x) Character. Most characters are generic. Airline safety cards use mostly female characters, while assembly and medical instructions use mostly male characters; custom characters are used very sparingly.
interference with the diagnostic process, we conducted a lab study to evaluate the effectiveness and legibility of our medical graphics. Specifically, we hoped to investigate if people’s comprehension and recall of diagnostic information and their response time improve given pictorial representations in contrast to pure textual information. We asked subjects to review materials digitally recorded in PHR using text and graphics about a simulated diagnosis taking place in a doctor’s office. To eliminate confounding factors like computer literacy, we used tabloid-sized paper booklets in the study instead of having participants operating on the computer. In addition, we provided the elderly participants with a magnifying glass so that they could read the information clearly.

6.1. Scenario development

With the help from the doctors in a local university medical center, we constructed diagnosis paths for four ailments: two upper respiratory infections (influenza and pneumonia) and two abdominal discomfort (gastroenteritis and appendicitis). For diagnosing each pair of ailments, the same questions are usually asked and the same physical tests are performed. However, the final diagnosis and treatments are different given different symptoms and test results. These stimuli reflect in real life how diagnostic examinations are performed, and require the participants to understand details and follow the whole process. We wrote a dialog script for each ailment (with and without the graphics), and constructed a questionnaire for each ailment to measure the participants’ comprehension and recall. Although some of the questions are different from ailment to ailment, materials and questionnaires were controlled for level of difficulty as judged by the doctors that we consulted.

6.2. Creation of visual vocabulary

We compiled a vocabulary of medical terminologies required in the diagnostic paths of the four ailments, including the common symptoms associated with the ailment, questions a doctor usually asks to identify the problem, the sets of tests (e.g. physical examination and lab tests) that are usually performed and the treatment usually provided. Since a health problem can be complicated by many factors, we focused on the most typical situations. We extracted 120 words and phrases from the diagnostic path, and picked the 46 terms that are the most critical according to the scripts to be visualized (Fig. 5).

6.3. Materials and experiment design

We took staged photos of simulated doctor–patient communication with actors/actresses recruited from a local university. Doctors from the medical center helped to script the dialogs of the complete diagnostic examination for each of the four ailments. Scenarios were laid out over multiple pages so that each page revealed a single fact and illustrative photo(s) about the diagnostic process (Fig. 6, top). A subsequent page contained diagnostic content related to the scenario, either in text form only (Fig. 6, middle) or supplemented with the medical graphics that we had designed (Fig. 6, bottom). This information showed what a doctor would record as the diagnostic process evolved.

Each participant was presented with four picture books, one for each ailment. Two contained diagnostic graphics; two were presented in text form. Since the diagnostic paths for the ailments in the same condition are quite similar, the stimuli were controlled for order. During the study, patients read each scenario, and then took a quiz (13 questions) that requires comprehension and recall of the diagnostic information presented in each picture book. Upon the completion of the entire study, we asked the participants to rate their perceived comprehension and recall of health information given graphical versus textual stimuli.

Examples of comprehension questions are:

Why does the doctor ask about lightheadedness?
- a. To determine if the patient has a fever.
- b. To determine the severity of the abdominal pain.
- c. To determine if the patient has any chronic disease.
- d. To determine if the patient needs more liquids.

What fluid does the doctor tell the patient to avoid, and why?
- a. Water, because it does not contain any nutrition.
- b. Broth, because it is too greasy and salty.
- c. Gatorade, because it is only for sports.
- d. Juice, because it irritates the stomach.

Examples of recall questions are:

What is the difference in the reading between standing-up and lying-down positions for pulse and blood pressure?
- a. The pulse decreases while blood pressure increases.
- b. The pulse increases while blood pressure decreases.
- c. Both the pulse and blood pressure decreases.
- d. Both the pulse and blood pressure increases.
- e. There is no difference.

What is NOT included in the doctor’s instruction?
- a. Avoid dehydration.
- b. Avoid starvation.
- c. Take medicine to relieve nausea.
- d. Take medicine to relieve diarrhea.

6.3.1. Participants

We tested the four scenarios with two populations that are more likely to have insufficient health literacy according to the
Developing Design Guidelines

Figure 6. Example page of the picture book about a scene of a clinic encountering (top), associated diagnostic template in text form (middle) and associated diagnostic template with graphics (bottom) in a PHR. In this example, the patient reported that she had a fever.

literature and our previous survey: the elderly and people from a different language and cultural background (U.S. Department of Health and Human Services, 2015). These two populations were reflected as two groups of participants, i.e. native English speakers of the age 65 and older (labeled as US-Old), and native Chinese speakers (non-native English speakers) of the age 20–55 (labeled as CN-Young). Native English speakers of the age 20–55 (labeled as US-Young) were used as the control group. We recruited eight participants for each group from the local community via mailing lists, posters and a participant pool.

All participants had college and higher level of education, and we tested their health literacy with three questions selected from the standard health literacy test: a medication warning label question (reading), a Body Mass Index Chart question (visual analysis), and a drug dosage question (calculation). Four out of the eight US-Young and three out of the eight CN-Young participants got all the health literacy questions right,
and none of the elderly participants had all the three questions answered correctly.

6.3.2. Measures
We gathered participants’ subjective feedback on the performance, interpretability, usefulness and liking of the medical graphics. In the post-study survey, the participants were asked to rank how well they understand and remember the diagnostic information discussed in the picture book given the graphical and textual template, respectively, on a seven-point Likert scale. Similarly, they were asked to assess the usability and usefulness of the graphics on a seven-point Likert scale. We also collected comments on experiences with the graphics.

Secondly, to evaluate the effectiveness of the medical graphics in comparison with text, we counted the number of correct responses to quiz questions for each participant group in different given different mode of stimuli. Each question is associated with one to four graphics, and we particularly examined the ones with big differences in accuracy between different representation modes or the ones that both modes performed poorly.

6.4. Data analysis
6.4.1. Perceived comprehension, recall, understandability and usefulness
Results from ANOVA showed that overall, people perceived to have better understanding ($F(1, 42) = 9.15, P = 0.004$) and memorization ($F(1, 42) = 10.49, P = 0.002$) of health information with graphics support than pure text. Such advantage of graphics over text is significant for the CN-Young group, and marginally significant for the US-Young group, but not for the US-Old group (Fig. 7).

In general, participants agreed that the medical graphics we designed were very easy to understand and very useful for following the entire diagnostic process (Table 4). In particular, all the Chinese participants commented that the graphics were extremely helpful when they encountered unknown medical terminologies. Six participants explicitly expressed strong liking of the medical graphics.

6.4.2. Reexaminations of design features
To gain a further understanding of how individual design features contribute to the interpretability and salience of medical diagnostic information, we examined individual responses to quiz questions. The findings were consistent with our predictions.

Features for highlighting critical facts and important location: when highlighting features were used, people with lower health literacy had better recall:

(i) Dates. In the graphics for appointment (Fig. 8, left), the current date and revisit date were marked in a different color, and all participants identified the next doctor’s appointment. On the contrary, the dosage graphics (Fig. 8, middle) do not highlight the days for the

Table 4. Means (Avg.) and standard deviation (SD) for the ratings (max. value = 7) on the perceived understandability and perceived usefulness in each participant group.

<table>
<thead>
<tr>
<th>Group</th>
<th>Understandability</th>
<th>Usefulness</th>
</tr>
</thead>
<tbody>
<tr>
<td>CN-Young</td>
<td>Avg. = 6.375, SD = 0.52</td>
<td>Avg. = 6.375, SD = 0.74</td>
</tr>
<tr>
<td>US-Old</td>
<td>Avg. = 6.250, SD = 0.89</td>
<td>Avg. = 6.375, SD = 0.74</td>
</tr>
<tr>
<td>US-Young</td>
<td>Avg. = 6.500, SD = 0.54</td>
<td>Avg. = 6.375, SD = 0.52</td>
</tr>
</tbody>
</table>


Figure 7. Perceived comprehension and recall for each of the three participant group, with maximum rating of 7 and standard errors.
prescription, and all groups performed poorly on the related question, which in reality may introduce risks of taking in medicine with the wrong dosage.

(ii) Location. The abdomen examination graphics (Fig. 8, right) allows the doctor to mark positions where the patient feels pain on the computer. This visual cue helped elderly participants to memorize the symptom (100% correct compared with 50% given text), which in practice may help them differentiate the ailment from other similar complaints. One participant commented that she could remember the picture as a whole instead of individual keywords for each location.

Features for reviewing essential details: zoom-in view and more realistic depiction improve people’s understanding of complex or ambiguous messages.

(i) Diagnosis. The graphics for pneumonia (Fig. 9, left) did not provide a zoom-in view for the inflamed air sacs, several people (3 CN-Young, 2 US-Old and 1 US-Young) mistook that the illness had something to do with bronchi, which may lead to mistreatment if people take home remedy. In contrast, the call-out view in the gastroenteritis helped people to map the diagnosis to its more common name ‘stomach flu’.

(ii) Examinations. The more realistic portrayal of the medical equipment and procedure assist participants with low medical literacy to draw connection to their existing knowledge, for example a CT scan (Fig. 9, right) is an imaging test, a lab test often involves blood test, and the doctor feeling the belly of the patient is performing a physical examination. With the help of the medical graphics, 75% of the CN-Young participants got the corresponding question right in comparison with none in the text condition. This will improve people’s understanding of the diagnostic procedure and allow them to seek proper tests when similar conditions occur.

Features for depicting cause–effect information: zoom-in views and more realistic depiction improve people’s understanding of complex or ambiguous messages.

(i) Medication. When the graphics for a medication depicts what symptom it is targeting (Fig. 10, left), people had better comprehension and recall when asked
Figure 10. Examples of multi-frame medical graphics with and without different visual cues: cough medicine (cause–effect ‘for’), shortness of breath when climbing stairs (situation ‘when’), and drinking normally and keeping the fluid down (sequence ‘then’).

about what and why medicine were prescribed. This may reduce situations where people take inappropriate medicines based on their personal belief.

One interesting finding was that people may interpret multi-frame graphics that depicts situations (X happens during Y) as a cause-and-effect message (X happens because of Y). For example, graphics for the shortness of breath when climbing stairs were interpreted as shortness of breath due to climbing. It suggests that different visual cues to connect the multiple frames should be used to increase the interpretability of the type of message for readers (Fig. 10). For instance, an arrow or a thought bubble can express causality; a call-out box with exploded view can present situations; while numbering can better indicate the chronological order.

Features for identification and demonstrating measurement results: numbers and scales in graphics reinforce the information and support quantification and comparison.

(i) Temperature. The graphics for taking oral temperature consists of a picture of patient having a thermometer in the mouth and another picture marking the results on a scale of normal, high and very high body temperatures. Both the US-Old (75% accuracy versus 50%) and US-Young (100 versus 50%) group had better recall of what kind of the fever the patient had when presented with graphical than textual diagnostic template (Fig. 11, left). We found similar effect on participants’ performance on the blood pressure recall question in Section 6.3 (Fig. 11, right).

In general, most of features proposed in our design guidelines successfully satisfy the criteria of visual saliency, legibility and interpretability.

6.5. Discussion

Overall, we received positive feedback on the understandability and usefulness of the medical graphics we designed from all participants, especially non-native English speakers from China. This implies the potential to help overcome the practical and motivational barriers of the broader adoption of PHR (Lober et al., 2006; Ratzan and Parker, 2000; Tang et al., 2006). We also identified several issues in the process. First, people coming from a country with different cultures, beliefs and social norms may have bigger disparities in healthcare practices than people living in the same neighborhood for a long time. Graphics designed according to an international standard may serve as a universal language that
Developing Design Guidelines

Figure 11. Example illustrations of measurement: temperature (two graphics on the left) and change in blood pressure under different posture (three graphics on the right).

helps bridge the gap. However, three Chinese participants suggested that they might resonate better with designs of regional characteristics.

Secondly, complex messages are particularly vulnerable to memory lapses Miller (1956), and indirect pictorial representations may fail to trigger a cued memory. A couple of people commented that complex graphics in addition to text may place greater demands on cognitive load. As a result, we observed that some of the elderly participants focused mostly on the dialogs, and skipped the page with the diagnostic template (text or graphics). This once again points out one of the great challenges in attracting attention to important diagnostic information, no matter how it is designed.

Thirdly, even for low-literate population, pictures should be used in conjunction with text, as pictures as a solo resource may not provide enough depth and unambiguity to convey complex medical information (Dowse and Ehlers, 2001). Two participants admitted that they did not pay attention to the differences between text and graphics. The interplay of text and pictures remains for further exploration.

7. CONCLUSION

The Quality of healthcare is critical to people’s wellness and well-being. Poor health literacy, or the ability to interpret and make judgments about one’s health, also affects well-being. In this paper, we explored how graphical medical information can assist in raising health literacy. We looked first at the domain of medical diagnostic information, used in PHRs. There is a lack of design principles about what features that could ensure the visual saliency, legibility and interpretability of this information, which aim to ensure the consistence among the design model, the system image and the user’s model (Norman, 1988). To this end, we conducted a two-stage investigation which confirmed that people have the ability to comprehend pictorial representations of medical conditions and treatments and improve their health literacy.

Our key contributions are 2-fold: (i) we provide a set of design guidelines for creating a medical visual vocabulary that can be universally understood; and (ii) we develop and test a digital graphical diagnostic template for PHR that employs medical graphics following the design guidelines we have generated. Our system can increase patients’ knowledge of the medical care services they receive, increase their familiarity with the information related to their personal health, and ultimately increase their interest and confidence in using a PHR.

In the future, we would like to apply the design of health-related visual information to the following applications: (i) potential self-diagnoses webpage; (ii) triage (getting information in ER); (iii) doctor–patient support program for remote areas; (iv) pediatrics; and (v) doctor–patient support for people with aphasia, a language disorder. Scenario 1 and 2 represent different contexts of obtaining diagnostic information. When people browse self-diagnoses webpages, they may be more relaxing than in an emergency room. Also, they acquire information mainly via reading rather than verbal conversations with healthcare professionals. Scenario 3–5 represent various populations who have great needs for medical care but are with relatively poor health literacy for different reasons. Our current lab studies involved a relatively small number of participants. We would like to test the design with more users, potentially in actual clinical practices to verify the efficacy, safety, and generality of our design implications, in comparison with existing systems. We will apply standard evaluations on the designed medical graphics, e.g. ISO/TC 145 Graphical symbols for appropriateness and public comprehension.

ACKNOWLEDGEMENTS

We thank Prof. Jodi Forlizzi, Prof. Marilyn Tremaine and Prof. Perry Cook for their insightful input, and thank the graphics designer Brian Mizrahi for the talented art work.

FUNDING

We thank NSF CIFellows grant #1019343 and the Princeton Kimberley and Frank H. Moss ‘71 Research Fund for sponsoring this work.
REFERENCES


Developing Design Guidelines


Rubin, E. (1958) Figure and Ground. New York, NY, Van Nostrand.


Wide, Range Inc. (1993) Wide Range Achievement Test (WRAT-3). Wilmington, DE.