Eyes on the Clinic: Accelerating Meaningful Interface Analysis through Unobtrusive Eye Tracking

Steven Rick[1,2], Alan Calvitti[2], Zia Agha[1,3], and Nadir Weibel[1,2].
[1]University of California San Diego, [2]VA San Diego Healthcare System, [3]West Health Institute, La Jolla
srick@ucsd.edu, alan.calvitti@va.gov, zagha@westhealth.org, weibel@ucsd.edu

***Abstract*—Electronic medical records (EMRs) are changing the way physicians work and how medical staff care for patients. While their widespread adoption promise many benefits and computationally powerful features for end users, they may also carry with them other unintended and troubling consequences. As part of a larger ongoing research study, we deployed an unobtrusive eye tracker in outpatient clinics to observe how physicians use their EMRs. We report on our experiences and we derive a methodology for successful eye tracking data collection in the clinic. Our results highlight multiple applications for the quantitative and qualitative assessment of EMR interfaces from eye tracking data collected in situ. We describe one of these applications, the association of eye movements with the specific task that physicians engage with in the EMR, and we discuss both next steps and future application of these results.**

Keywords*—*Eye-tracking, observational research, medical office, Electronic Medical Records, user interface research

I. Introduction

Electronic medical record (EMR) systems are evolving to offer powerful new features for data retrieval and diagnosis as well as computationally sophisticated algorithms to mine patient data and alert physicians to potential concerns. But these growing health IT systems may also carry with them unintended and troubling consequences due to the way they are being designed. These systems have begun to be examined closer as their usage has become more and more pervasive, and poor system design in conjunction with improper use has led to rising concern around EMR systems potentially endangering patient safety and decreasing quality of care [1]. Just as more research needs to be done to understand the consequences of electronic medical system adoption, there is a need for methods and tools that enable their critical assessment to inform developers and improve health IT for the future.

User research often poses many challenges, and these are further complicated in the health domain. HIPAA (Health Insurance Portability and Accountability Act)[1] and other health-data regulations, in place for the protection of patients and their privacy, create great overhead for researchers intending to assess systems rather than the actual data within them. Due to these concerns, EMR system research often takes place in patient simulations. Data gathered in that manner is often not a true representation of how the systems are actually being used in real and dynamic environments [9]. For this reason, we stress the importance of in situ data collection from real physicians while they hold actual visits with real patients.

This work is a part of a larger ongoing research study quantifying EMR usability to improve clinical work. It aims to build quantifiable profiles of EMR use within clinical settings through the application of easily deployable sensor technologies in real-world settings [10]. By combining these data with clinical context, we aim to better understand the consequences of how EMR use affects the communication and work flow of physicians. In this paper we focus solely on the role of eye tracking in this setting and we outline our approaches for understanding EMR interactions through unobtrusive eye tracking data collection. We also discuss our data analysis methods which resulted in new insights that show potential to empower current and future research on health IT.

II. Eye Tracking in the Clinic

Classic usability methods based on think-aloud protocols, task analysis, etc. require researchers to be present while the activity occurs. While these methods are useful, in real clinical environments the presence of research staff in the room while patient and physician meet may hinder the natural activity patterns that would have occurred had the environment been more private. For this reason we emphasize a remote sensing approach to data capture where patient and physician are left to work together without external influence from research staff.

Eye tracking is a technology that allows for insight into the conscious and even subconscious activity of users. More specifically, eye-movement can provide useful insights for in-depth analysis of the usability of interfaces [7]. Although often seen as particularly challenging for eye tracking data collection, recent work has started to use mobile eye tracking glasses in the medical office [8]. Researchers collected data from eye tracking glasses and used contextual inquiry to gather additional insights about the information that the medical providers were accessing and the reasons behind it. While insightful, worn systems like eye tracking glasses present the potential to disrupt and interfere with natural behavior patterns, as those devices are not regularly worn by the medical staff.

[1]http://hhs.gov/ocr/privacy

PervasiveHealth 2015, May 20-23, Istanbul, Turkey

DOI 10.4108/icst.pervasivehealth.2015.259276

This could unintentionally alter the dynamic between staff and their patients. We chose to emphasize an unobtrusive approach to eye tracking so that our observation of physicians in their natural environment was not impacted by the data collection. A small remote (non-worn) eye tracking sensor allowed for the technology to disappear into the background of the clinical encounter, something that worn systems cannot do.

Nielsona et al. ran another study looking at emergency physicians and their use of EMR systems. They used a Tobii T60 monitor with built in eye tracking to less intrusively track visual attention of the physicians on screen [6]. They aimed to quantify how laboratory results were accessed across a number of physicians, but chose to pause data recording while the physicians were seeing patients.

Our approach expands those first attempts to collect naturalistic data in the clinic by capturing entire patient-physician encounters, from arrival to departure. In this way we intend to gather information key to helping us understand the complete flow of events over the course of the entire encounter.

III. Research Environment

To track visual attention and eye activity in the clinic, and specifically to be able to investigate how eye tracking is associated with the usage of EMRs, we leverage a remote eye tracking device, the SMI RED-m,[2] and we integrate it within an extensive infrastructure for multi-modal data collection [10]. The small device is configured to collect data at a rate of 120 Hz, and is mounted unobtrusively beneath the primary monitor of the physician's desktop computer as shown in Fig. 1. Once set up, the device's positioning is adjusted and the physician's eyes are calibrated. Upon successful calibration (vertical/horizontal offset of 1°of visual angle or less) the system is left to run without further intervention for subsequent patient visits. By running in tandem with an Epiphan DVI2USB3.0 frame grabber and the SMI Experiment Center software, we capture a rich set of data providing detail about both where each of our participant's visual attention is on screen as well as what they are currently viewing providing context behind the tracking.

We have deployed the system in three different Veteran Affair (VA) outpatient clinics in San Diego (California, USA), where we have captured data from 12 physicians across 88 patient visits, which averaged 30 minutes in length, for a total of roughly 2,500 minutes (40+ hours). We collected data from one physician each day while they worked on a single PC over the course of multiple back-to-back patient visits. In every clinic, the rooms were oriented such that the patient sat with their back to the wall, facing their physician, while the physicians sat at a desk next to the patient facing the opposite direction towards the computer monitor (see Fig. 1).We requested that the physicians try to not move their monitor around – and as such not move the eye tracker from its calibrated position – but they were otherwise left to work however they wish. The nature of this sort of uncontrolled data collection gave us great insight into the natural work flow of our participants, but also introduced a reasonable amount of noise in the data that had to be overcome.

[2]http://www.smivision.com

IV. Setting Eyes on the Clinic

Three procedures showed a major impact in terms of data quality among our 88 eye tracking sessions in the clinic and resulted in successful data collection, namely *calibration*, *data cleaning* and *contextual coding*.

A. Calibration

During the pilot phase of our study (patients 1-6), we realized that when performing eye calibration our participants would sit up straight, hold their heads quite still, and calibrate very well. However, once they began their work and forgot about the eye tracking, they quickly slouched and repositioned themselves into a more comfortable posture, disrupting the sensor's calibration. This caused four out of our six pilot recordings to report no eye activity data. With calibration the only point during our data collection that required the researchers to be present, we needed a remedy to be introduced prior to calibration. We learned that by having each physician work with their computer for a few minutes before any calibration took place, each participant became engaged in their regular routine and moved into their natural computer use posture. Emphasizing the lack of a need to sit up straight, or hold their head still, improved the quality of eye tracking after calibration. We no longer experienced problems due to calibration issues after the conclusion of our pilot.

B. Data Cleaning

In the natural setting of the clinic, eye tracking data was noisy. Our subjects were free to move around in their environments, work with paper documents and other interfaces with their patients. This variance in behaviors across physicians as well as with the same physician across multiple patients introduced differences in eye tracking signals: some just tracked better while others were more erratic. To compensate for these larger scale differences, as well as the fast nature of the eye, our next goal focused on finding a method to help clean up and normalize our data. We developed two methods to assist with cleaning eye tracking data: *down-sampling* and *continuous attentional windows*.

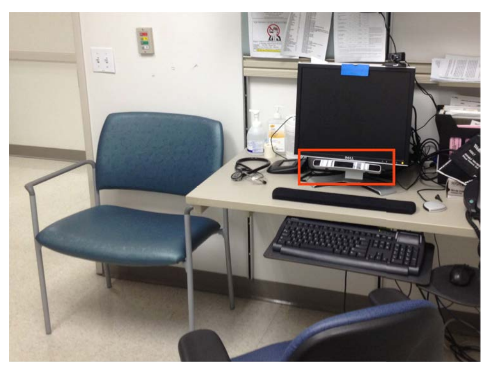

Fig. 1. Medical exam room seating arrangement seen from the exam table. Note the location of the patient chair on the left of the image: when the patient sits, he/she faces the physician sitting and typing at the computer on the right. The eye tracker mounted beneath the monitor is highlighted in red.

Down-sampling – We found that much of the information we capture was spatially redundant. The high sample rate of our sensor captured jitter due to momentary errors in tracking and the physiology of the eye. By reducing our stream to every fourth data point, we found that the same quality was available at a fraction of the cost, both in file size and computational demands. This reduction carried enough continuous information to be able to accomplish analysis of attention, scanning patterns, and other behaviors while using a fraction of the computation time. This accelerated our usage of the data and allowed for quicker turnaround of result intermediaries thanks to only needing to analyze an average of 15,000 lines of data instead of 60,000 per patient encounter.

Continuous attentional windows – Even after data was reduced, there was still a range of momentary drops in the tracking data associated with blinking, looking down at the keyboard while typing, physical movement, eyelash interference, and other anomalies. We decided that these very brief drops in the middle of continuous data needed to be differentiated from longer dropouts associated with looking away from the monitor or doing other tasks not related to computer usage. We developed a method similar to [5] for windowing the tracking data into segments. We take into consideration that data temporarily missing for 400ms or less should account for most natural interferences, such as blinking, and other oculomotor activities. Thus we are able to combine multiple segments of data in close temporal proximity, regardless of gaps in tracking, to create windows of data for analysis.

C. Contextual Coding

Lastly, by exploiting sensor data fusion, as enabled by our infrastructure [10], eye activity data could be aligned alongside other streams. To aid in our endeavor to better understand EMR system use and interface design through eye tracking, we therefore merged the eye movement data with the context of the activity performed on the computer and connected eye movements with patient care activities. In particular, we segmented our eye data based on what current EMR function was used by the physician. Manually review of the screen recordings from each patient visit let us time code the individual EMR functions accessed and clinical tasks performed and this allowed us to segment eye tracking based on clinical activities such as information retrieval from lab results, documentation in clinical notes, ordering medications, and more from within the EMR system.

V. Behavioral Assessment of Eye Activity

To further our understanding of eye activity dynamics within EMR use in the clinic, it was important to develop a method for classifying the types of behaviors that physicians were engaging in while working with their computers. From manual review of our data, we noted that much of the behaviors performed by physicians could be categorized into three groups: 1) searching, 2) focusing, and 3) reading. This insight drove the development of a preliminary behavior pattern algorithm to categorize our windows of coded eye activity data.

To align with the three distinct types of activities, we developed a method to classify eye tracking windows as *jumpy*, *linear*, or *focused* by setting two thresholds. First, in order for data to be classified as linear, the angle generated between two sequential data points must not deviate more than 5 degrees from the plane established by the previous two sequential data points. Next, if data did not qualify as linear, it was checked to see if it could be classified as focused. As discussed in [11], fixations that occur in close proximity to one another (within 64 pixels) are considered recurrent. In our filtering algorithm we relaxed the threshold and settled upon a 70 pixels radius between points for qualifying data as focused. Finally, if a data point did not qualify as linear or focused it was considered jumpy since it was neither moving along the same plane of motion or in close proximity to the prior point.

We ran our classification algorithm for each data point within each 'window' of our eye tracking data, and assigned a point to one of the three categories (focused, linear or jumpy). Once a segment was scored, the values were averaged and if a single category held over 70% of the weight, the window of data was considered predominantly showing behavior from that category. This preliminary categorization method enabled further analysis such as how a physician's eye activity behavior changed over time compared to the EMR function or clinical task they were performing. Across our data set, we analyzed how behavior differed between Primary Care physicians and Specialists (Fig. 2), and how it differed between different types of EMR tabs across physicians (Fig. 3).

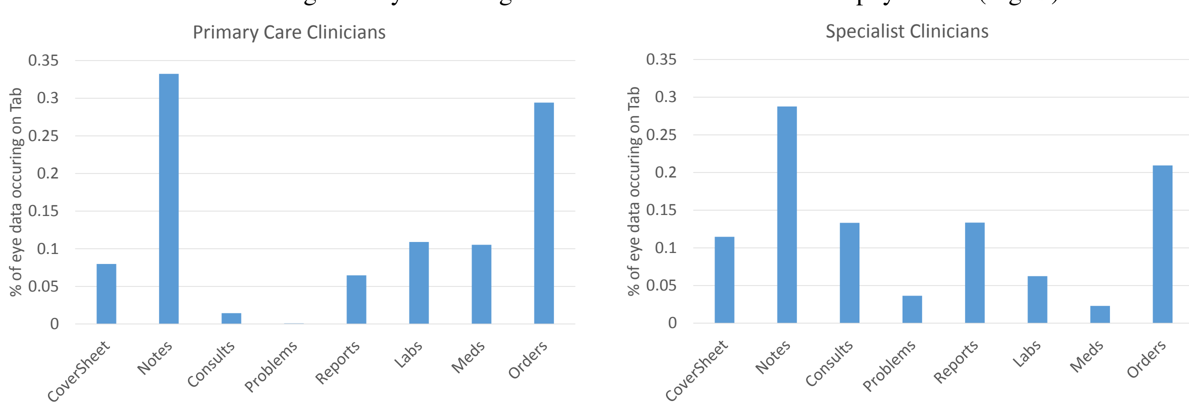

Fig. 2. Difference in distribution of gaze activity between *Primary Care* and *Specialty* clinicians. The behavior of the two kind of physicians are largely similar, with clinicians dedicating a similar amount of gaze activity across a number of EMR tabs/tasks. Both predominately focused on activity in Meds and Orders.

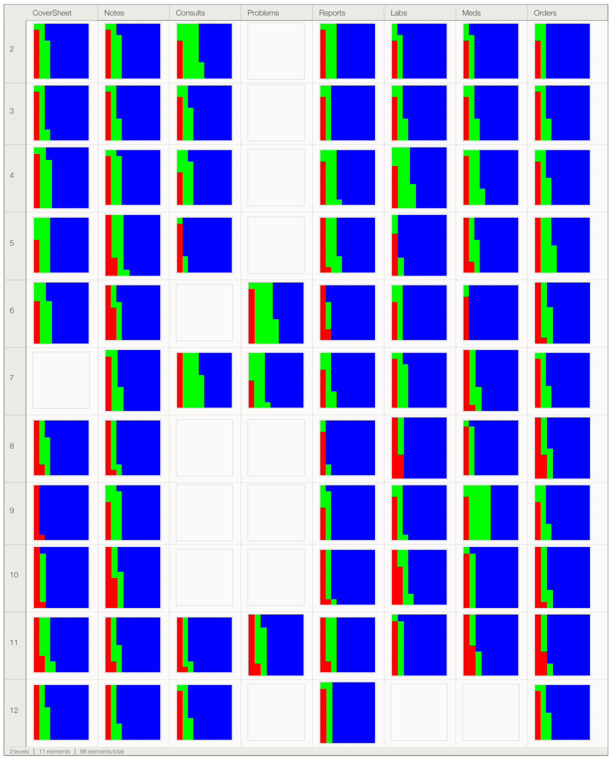

Fig. 3. Behavior distribution across physicians and tabs. The predominant behavior is jumpy (blue), but occasionally on Notes, Orders, or Labs there is a higher prevalence of focused (green) or linear (red) behavior.

VI. Discussion and Future Work

The work presented is ongoing and seeks to produce more refined and complex analytic methods of usability assessment around EMR usage. As discussed by [3], medical practices are embedded within the usage of complex data and documents, and today's EMR systems are the gatekeepers standing between physicians and the data of their patients. If these systems are designed well and meet the needs of doctors, physicians will be given the ability to work efficiently and improve the quality of clinical care that they deliver. Inversely, if these systems are not designed well we will very likely notice a surplus of energy being devoted to the task of searching through and retrieving important information. As our behavior classifier started to show, the eye activity of our 12 physicians while they engaged with their EMR system across 88 patients was predominated by searching behavior. More work must be done to understand this behavior in depth, but initial results indicate that the EMR system observed may not be organized in a manner that is conducive to the way doctors work.

In terms of our specific research in the real-world medical office, we are planning to exploit these techniques to assess, compare, and contrast how two different large scale EMR systems are used by physicians working in two large medical institutions. Employing eye tracking in a naturalistic and ecologically valid environment will allow us to uncover how the design of these systems may cause common behaviors and highlight usability problems across different health IT systems.

This kind of data opens up a range of new and exciting possibilities that might heavily impact eye tracking research outside of the lab. Collecting high-quality data will enable the use of cutting-edge computer vision approaches to speed up segmentation and annotation as compared to our current manual procedure which is very time consuming. Feature detection algorithms can be run on the screen-capture videos to automatically tag specific tasks and interactions. The different components and features of an EMR interface could be modeled to accelerate task segmentation, similar to how [2] used color and edge detection to segment components of an image. Finally, eye motion traces could be modeled to build user activity classification systems that can segment activities from one another, similar to how [4] differentiated between reading, searching, or memorizing a scene.

Acknowledgments

This research was funded by AHRQ grant 1 R01 HS021290-01A1, Zia Agha PI. We would like to thank all Co-Investigator and collaborators on the QUICK study who have been involved in our eye-tracking study.

References

[1] S. Bowman, "Impact of electronic health record systems on information integrity: Quality and safety implications," *Perspect Health Inf Manag*, vol. 10, 2013.

[2] J. Fan, D. Yau, A. Elmagarmid, and W. Aref, "Automatic image segmentation by integrating color-edge extraction and seeded region growing," *IEEE Trans. Image Process*, vol. 10, no. 10, pp. 1454–1466, 2001.

[3] C. Heath and P. Luff, "Documents and professional practice: bad organisational reasons for good clinical records," in *Proc. CSCW '96*, 1996, pp. 354–363.

[4] J. M. Henderson, S. V. Shinkareva, J. Wang, S. G. Luke, and J. Olejarczyk, "Predicting cognitive state from eye movements," *PLoS ONE*, vol. 8, no. 5, p. e64937, May 2013. [Online]. Available: http://dx.doi.org/10.1371/journal.pone.0064937

[5] M. Johns, K. Crowley, R. Chapman, A. Tucker, and C. Hocking, "The effect of blinks and saccadic eye movements on visual reaction times," *Attention, Perception & Psychophysics*, vol. 71, no. 4, pp. 783–788, 2009.

[6] J. A. Nielsona, R. N. Mamidala, and J. Khan, "In-situ eye-tracking of emergency physician result review," *Stud Health Technol Inform*, vol. 192, p. 1156, 2013.

[7] A. Poole and L. J. Ball, "Eye tracking in human-computer interaction and usability research: Current status and future," in *Prospects?, Chapter in C. Ghaoui (Ed.): Encyclopedia of Human-Computer Interaction. Pennsylvania: Idea Group, Inc*, 2005.

[8] N. Segall, S. Dunbar, E. W. Moretti, R. A. Schroeder, J. M. Taekman, and M. C. Wright, "Contextual inquiry to support the design of electronic health records," in *Abstracts of the 2013 ASA Meeting*, 2013.

[9] B. W. Tatler, M. M. Hayhoe, M. F. Land, and D. H. Ballard, "Eye guidance in natural vision: Reinterpreting salience," *Journal of Vision*, vol. 11, no. 5, p. 5, 2011.

[10] N. Weibel, S. Rick, C. Emmenegger, S. Ashfaq, A. Calvitti, and Z. Agha, "Lab-in-a-box: semi-automatic tracking of activity in the medical office," *Personal and Ubiquitous Computing*, pp. 1–18, 2014.

[11] D. W.-L. Wu, N. C. Anderson, W. F. Bischof, and A. Kingstone, "Temporal dynamics of eye movements are related to differences in scene complexity and clutter," *Journal of vision*, vol. 14, no. 9, p. 8, 2014.