Transmitting Narrative: An Interactive Shift-Summarization Tool for Improving Nurse Communication

Angus Forbes[*]**, Mihai Surdeanu**[†]**, Peter Jansen**[‡]
School of Information, University of Arizona

Jane Carrington[§]
Department of Nursing, University of Arizona

ABSTRACT

This paper describes an ongoing visualization project that aims to improve nurse communication. In particular, we investigate the transmission of information that is related to potentially life-threatening clinical events. Currently these events may remain unnoticed or are misinterpreted by nurses, or most unfortunately, are simply not communicated clearly between nurses during a shift change, leading in some cases to catastrophic results. Our visualization system is based on a novel application of machine learning and natural language processing algorithms. Results are presented in the form of an interactive shift-summarization tool which augments existing Electronic Health Records (EHRs). This tool provides a high level overview of the patient's health that is generated through an analysis of heterogeneous data: verbal summarizations describing the patient's health provided by the nurse in charge of the patient, the various monitored vital signs of the patient, and historical information of patients that had unexpected adverse reactions that were not foreseen by the receiving nurse *despite* being indicated by the responding nurse. In this paper, we introduce the urgent need for such a tool, describe the various components of our heterogeneous data analysis system, and present proposed enhancements to EHRs via the shift-summarization tool. This interactive, visual tool clearly indicates potential clinical events generated by our automated inferencing system; lets a nurse quickly verify the likelihood of these events; provides a mechanism for annotating the generated events; and finally, makes it easy for a nurse to navigate the temporal aspects of patient data collected during a shift. This temporal data can then be used to interactively articulate a narrative that more effectively transmits pertinent data to other nurses.

Author Keywords

EHRs; text analytics; interactive verification; temporal data; heterogeneous data alignment; health informatics.

ACM Classification Keywords

H.5.m. Information Interfaces and Presentation (e.g. HCI): Miscellaneous

INTRODUCTION

The introduction and subsequent adoption of the Electronic Health Record (hereafter, EHR) is a step forward in automating data collection useful in the analysis of patient health and

[*]angus.forbes@sista.arizona.edu
[†]mihais@sista.arizona.edu
[‡]upontheturtlesback@gmail.com
[§]jcarrington@nursing.arizona.edu

prevention of adverse health events [9]. However, unintended consequences mitigate the effectiveness of the EHR [8, 16, 27]. Most problematically, nurses alternately taking care of a patient are unable to effectively sift through the large amount of data available via the EHR in order to find pertinent information. Studies of nurse behavior has found that many nurses make an effort to talk to each other face-to-face as they change shifts [6]. Ideally, these face-to-face conversations provide a way for the *responding* nurse (leaving the current shift) to explain his or her interpretation of the patient's well being to the *receiving* nurse (starting the next shift). The absence of this dialog may account for an alarming number of miscommunications that have lead to catastrophic events in patient health [27]. For instance, a report by the Institute of Medicine [18] finds that up to 98,000 patients die per year as a result of complications of therapy due to ineffective communication and, moreover, that errors in communication cost US hospitals an estimated $12 billion annually [13]. Although providing verbal summarizations of patient health during in the hand-off between responding and receiving nurses can be helpful, these summarizations can themselves be misinterpreted or ignored [7].

In this paper, we describe an interactive visualization system that augments EHRs to improve nurse-to-nurse communication. Our system uses a novel application of machine learning and natural language processing techniques to generate a series of potential *clinical events* [20] and, furthermore, can offer reasons why these events are plausible based on an analysis of: the vital signs recorded in the EHR; the verbal "hand-off" summary of patient health made by the responding nurse; and historical EHR data of patients that had unexpected adverse clinical events. In particular we look at data related to six categories of clinical events that are most likely to be a precursor to unexpected patient death: uncontrolled pain, sudden fever, bleeding, changes in respiratory status, changes in level of consciousness, and changes in output. We further describe a novel, interactive shift-summarization visualization tool that provides: 1) the automatic proposal and notification of these potentially adverse clinical events; 2) a way for nurses to verify the likelihood of these events; 3) a mechanism for annotating the automatic proposals and for allowing the nurse to propose their own events; 4) and finally, an interactive tool that lets the responding nurse associate vital signs as they evolve over time with an overview narrative, making it easier to indicate highly pertinent data to the receiving nurse. In effect, our system aims to replicate, to some extent, aspects of face-to-face nurse communication that may have been lost through the introduction of EHRs, and to formalize the verbal summarization that some nurses have improvised. That

is, our visualization system aims to make it easy for nurses to create, transmit, and verify narratives about patient health that augment their expert knowledge for improved decision making.

Both the development of a computational model to predict clinical events and the creation of interactive visualization tools based upon this model make use of data gathered from real-world scenarios where EHRs are used by nurses in making decisions regarding patient treatment. Currently, we have extensively annotated data generated through interviews with 37 nurses using EHRs who oversaw patients who died unexpectedly, and we are streamlining the interviewing process in order to gather 100 more samples of nurse (mis)communication [14]. Although this work is ongoing, it brings together interdisciplinary expertise in health informatics, machine learning, and text analytics with the aim of creating an effective interactive visualization tool with the potential to save lives.

GENERATING AND EXPLAINING CLINICAL EVENTS

The computational modeling component of our visualization is divided into four interrelated modules, each making use of information available in the EHR and/or the transcriptions of verbal summarizations, or "hand-off" reports, created by the responding nurse. The first two models provide a detection facility for clinical events, as well as a probabilistic outcome prediction mechanism to determine the relative likelihood of an undesired outcome given the detection of a clinical event. To make certain that the hand-off report is properly connected with the EHR data, a third model grounds the narrative of the hand-off report with vital sign measurements in the EHR. These links between the two content modalities are further exploited in the fourth model, which generates an automated summary of the hand-off report, highlighting information that is predictive of a negative patient outcome.

Previous research, such as [12, 17], explores the usefulness of using natural language processing techniques in the context of health informatics and medical decision making. Our system introduces the idea of *aligning* the EHR data that tracks vital health signals from EHR with the verbal hand-off reports created by the nurses. This central feature of our system stems from the observation that there are clear connections between the two data sets, and that the responding nurse typically explains the most relevant aspects of the EHR in the hand-off report. The information in these two modalities therefore is jointly extracted, such that in the case where a vital sign measurement is highlighted in the EHR, the corresponding text in the hand-off report that describes and elaborates on this vital sign should also be extracted to provide the receiving nurse with more contextually useful diagnostic information.

Grounding the text of the hand-off report in the vital sign measurements of the EHR can be viewed as an alignment task that can be formalized as a statistical machine translation (SMT) problem. Conceptually, the EHR and the hand-off report can be viewed as as two documents describing the same information in different "languages." The alignment between these two languages (which grounds the hand-off report in terms of the EHR data) can be learned using SMT algorithms

(e.g., [21]). While the SMT analogy is strong, there are clear differences between SMT and the task at hand. Most importantly, most of the information in the EHR is stored as continuous numeric values (i.e., vital sign measurements), which cannot be directly used in an alignment model. Using the same SMT analogy, this is equivalent to a language with an infinite vocabulary (one "word" for each numeric value of each vital sign). To mitigate this problem, we discretize the numeric measurements using a three-point scale: below normal, normal, above normal. Because the hand-off reports are considerably more verbose than the EHR, the granularity of the information stored in the two documents varies considerably. For example, an nurse will often choose to explain a single vital sign measurement using one or more sentences. This is different than SMT, where individual words (or small phrases) in the source language are aligned to similar words (or phrases) in the destination language. Fortunately, methods for dealing with this disparity are common in other fields where SMT models are applied to non-translation tasks, such as question answering [28], where one of the "languages" is considerably more verbose. Our alignment model is trained using matched pairs of EHRs and hand-off reports for a given patient. While SMT models typically require a large set of aligned texts in each of the two languages to arrive at seamless translations, restricting our SMT model to domain-specific alignments allows us to train using a relatively small corpus.

Our models produce the following main results (listed below), which are then displayed or are interactively investigated via the shift-summarization tool. The evaluation criteria for the computational models are a) the success at detecting events, and b) the accuracy in predicting patient outcomes. Standard evaluation of these metrics consists of comparing accuracy (in terms of sensitivity and specificity) against a gold-standard set of data created by expert human annotators, which we are currently in the process of generating. Our inference justification model will likewise be evaluated against alignments generated by human experts.

Clinical Event Detection Facility

Given an electronic health record for a patient, a nurse-to-nurse verbal hand-off transcript, or both, the computational model serves as a multiclass classifier able to detect whether a clinical event has occurred or not. Moreover, the model provides the central predictive features that it used to arrive at this inference.

Outcome Prediction

Given an electronic health record for a patient, a nurse-to-nurse verbal hand-off transcript, or both, the computational model generates a probability distribution over a set of possible patient outcomes (neutral, extended length of stay, failure to rescue). This model also provides the central predictive features used to arrive at this inference.

Inference Justification

Given a detection or prediction inference, the computational model is able to identify specific features in the EHR or hand-off report that it used to arrive at that inference. Using the grounding of the natural language of the verbal report in EHR

data, the model identifies which EHR features align with particular sections of the verbal report – providing justification from multiple sources – or shows where information contained in either the EHR or verbal report was not detected in the other modality, suggesting a possible error where patient information was miscommunicated, incorrectly recorded, or lost.

EHR and Hand-off Report Summarization
Given a hand-off report and corresponding EHR, the computational model generates an automated summary highlighting information that is highly predictive of a negative outcome, based on an analysis of historical patient data indicating particular patterns that led to adverse events.
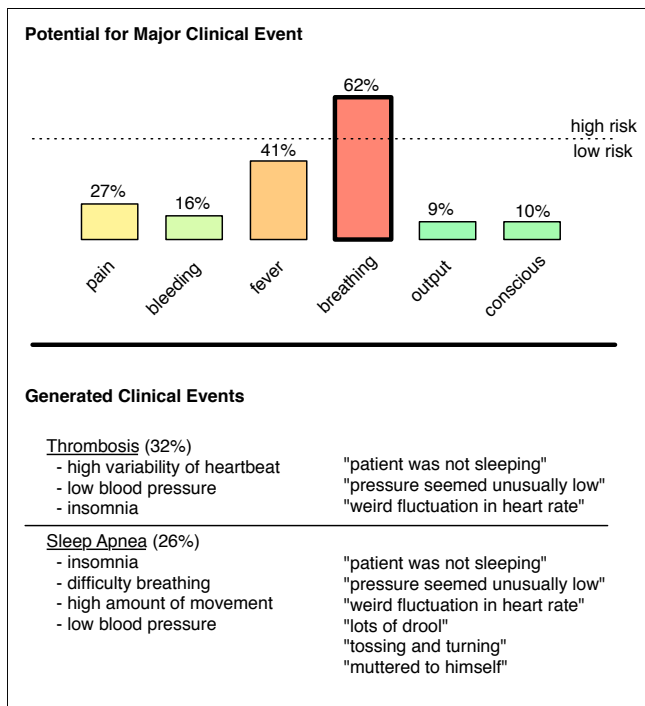

Figure 1. A prototype of the overview page of our shift-summarization tool. At the top we see a simple bar chart indicating high-risk clinical events (determined through an evaluation of historical EHR data). At the bottom we see a list of events generated by our system (based on expert rules), alongside related textual snippets from the nurse's "hand-off" report.

VISUALIZATION TOOLS FOR EHRS
The existing design of EHRs is graphically cluttered and parsing patient data can be cognitively-taxing, even for experts [30]. Moreover, in the attempt to provide information about all possible indicators of clinical events, ironically, the most relevant and potentially life-threatening of these can be obfuscated [7]. While previous work has explored different approaches to visualizing and managing the complexity of EHRs, for example [3, 22], our shift-summarization tool provides enhancements to EHRs to improve nurse comprehension of and communication about patient data that is related in particular to signals of a sudden and unexpected change in patient condition. The primary components of our visualization enhancements are based directly on the data generated by

the computational models: the alignment between EHR data and the responding nurses verbal summarization of the events during the shift; a series of generated explanations as to why these events happened; and an overview likelihood, based on historical data, of whether or not these generated events could lead to adverse changes in patient condition. Additionally, we incorporate temporal aspects of the EHR data in order to augment the receiving nurses decision making when verifying the recommendations produced by the data analysis module.

Specifically, our visualization enhancements provide: a way to *notify* nurses about high-risk situations; detailed information that nurses can use to *verify* these notifications; a way for nurses to meaningfully *annotate* their interpretation of the patient health in relation to inferred clinical events; and finally, a system that encourages a nurse to link annotations to temporal data in order to *narrate* the overall story of how the patients health evolved over the course of a shift. Each enhancement builds on the previous, adding functionality that increases the effective transfer of pertinent information between nurses when using EHRs. We are implementing the prototype of our shift summarization tool on an iPad tablet, but expect also to port it to desktop environments or other mobile devices.

Enhancement 1: Shift Summarization (Notify and verify)
Our primary enhancement clearly presents the results of our computational analysis. All six of the major clinical events are listed, and high-likelihood events (as determined via a historical analysis of the EHRs of patients who suffered these events) are highlighted. The receiving nurse is thus, at a glance, informed as to whether or not the patient is in imminent danger.

In addition to this high-level overview specific to the major clinical events, we also generate a list of other inferred clinical events that occurred during the shift and provide the evidence as to why they were inferred. That is, for each of the generated clinical events, we link to the EHR monitoring data or the transcription of the nurse's verbal summarization that led our system to conclude that the event occurred.

Explicitly presenting data that shows the logic behind the computational inference functions both as a way to create trust in the system, as has been shown in research on (or implemented in) recommendation systems, such as [10, 11, 15, 23, 26], and also provides a starting point for verifying or invalidating the automatic notifications, which could be especially important in the case of false alarms. A third component of this first enhancement is to provide an interface for browsing the alignment between EHR data and text. This enables the nurses to verify results of our automated system easily, and also to search freely for patient-specific information that could bolster or invalidate the automated recommendations. Figure 1 shows a prototype of this enhancement. At the top we see a simple bar chart indicating high-risk clinical events. At the bottom we see a list of events generated by our system, alongside related textual snippets from the nurse's hand-off report. Figure 2 shows an example of how a nurse selecting a particular explanation for a generated clinical event can instantly see more detail about the aligned text

and EHR vital signs that lead our system to present this explanation.

Enhancement 2: Issue Tracking the EHR (Annotate)
This enhancement aims to promote dialog between the nurses, to foster engagement with relevant information about patient health, and to provide accountability for the nurses caring for the patient.

The generated clinical events are essentially an *interpretation* of the raw data and the nurses verbal summary of the shift. Moreover, the alignment process between the EHR data and the verbal summary is also based on encoded *assumptions*. We provide a system for the nurses to agree or disagree with the automatically generated clinical events and to annotate them with additional pertinent information. The annotations are in the form of a predefined comment – such as "agree", "disagree", or "inconclusive" – and, optionally, space is available for further detailed commentary. By providing a mechanism that operates, essentially, as an issue-tracker, nurses have an opportunity to create and respond to forums about particular events that may be important to the patients health. In issue-tracking software or websites used for software projects (such as [1, 2]), this type of commentary is used to build consensus on interpretation, to expedite decision making, and to facilitate conversation [5]. By requiring the nurses to annotate each of the generated events as well providing the ability to define their own, we instantly create a focused dialog about the relevant issues regarding the health of the patient. Furthermore, this mechanism creates a trail of accountability, as the nurse can explicitly explain their reasons for disregarding an event, or modify the reasoning behind why the event occurred.

The enhancement is split into two related components, one for the responding nurse and the other for the receiving nurse. First, we allow the responding nurse to evaluate the shift summarization created by Enhancement 1. In particular, we allow the nurse to annotate the generated events, either with predefined terms, or with a more detailed textual explanation. Thus the receiving nurse has more information regarding consensus between the automated interpretation and the nurses interpretation of events in the shift. Second, we similarly provide a space to indicate agreement or disagreement with generated events and allow the receiving nurse to indicate that he or she has read and the responding nurses annotations as well as space to provide additional commentary.

Enhancement 3: Telling a story via temporal data (Narrate)
This enhancement extends the annotation mechanisms described in Enhancement 2, allowing the responding nurse to link their interpretation of the patient's health to particular events in time. That is, we allow the responding nurse to create a *curated timeline* of the patient's health as it evolved over the course of the previous shift. The receiving nurse can then use this temporally-contextualized data to augment his or her decision making process during the current shift. Previous research has investigated visual information seeking over temporal data across multiple EHRs [4, 24, 29, 30]. Our project emphasizes the temporal aspects of the patient data over the course of a single shift. This enhancement allows

the responding nurse to browse the generated events and annotations along with the temporal EHR data (e.g. the "flow sheet") and then to create a visual narrative of how the nurse reasons about the possibility of clinical events. Narratives are made up of pertinent, sequential events, and providing a system that allows the nurse to explain the events of the shift in a narrative format attempts to replace an important aspect of face-to-face communication that is otherwise lost in the hand-off. Figure 3 shows a sketch of the proposed enhancement in which an annotated timeline is associated with potential clinical events. The circles indicate highlighted raw vital signs from an EHR flow sheet that led the responding nurse to make a comment.

Figure 2. An example of how a nurse selecting a particular explanation for a generated clinical event can instantly see more detail about the aligned text and EHR vital signs that lead our system to present this explanation.

Evaluation Methods and Considerations
Our shift-summarization tool aims to clearly represent the automated interpretation of patient data as well as the sentiment of the responding nurse and, second, to ensure that the receiving nurses trust the representation sufficiently to incorporate it into their decision making. Our system is meant to *augment* (and not replace) nurses' expert skills, and our evaluation is aimed not only at justifying appropriate visualization methods but also to measure the effectiveness of their integration into current practices.
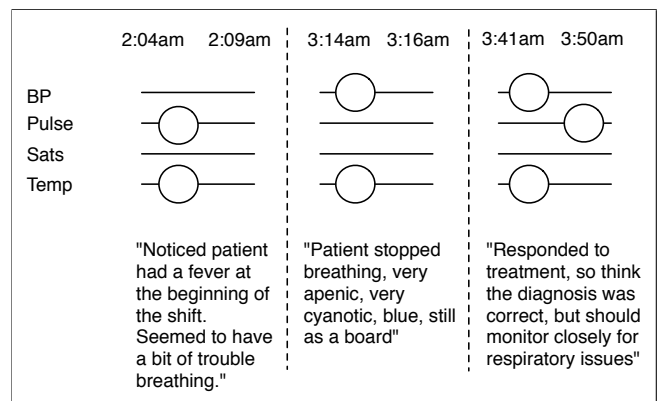

Figure 3. An annotated timeline that helps to *tell a story* by relating pertinent points of data on particular vital sign streams to nurse interpretation.

Our two main contributions throughout these enhancements are a) clearly representing the results generated by our computational models, and b) providing an interactive interface with which to support rapid exploration of EHR data. We hypothesize that each additional enhancement provides increasingly more effective communication leading to more accurate

diagnoses of at-risk patients. We further hypothesize that our system increases the nurses ability reason about salient information, while simultaneously reducing the amount of time they spend filtering out irrelevant data. To test these hypotheses we plan to run user-studies with nursing students and practicing nurses. These user studies will examine whether an expert using the proposed enhancements performs better than the automated system alone. We are also concerned about the level of trust an expert might have in the automated system and also with the amount of time it takes to use the enhancements. Obviously if there is no trust in the system or if it causes an undue burden, then nurses will be less likely to use the visualization tools in real-world situations. Thus, in addition to the user-studies, we will also conduct cognitive walkthroughs in order ascertain the ease of use and potential rate of adoption of the tools. As discussed in [19] and [25], the iterative design of information visualization tools based on detailed feedback from cognitive walkthroughs with domain experts can be an effective way to create systems that will be useful in real-world environments.

CONCLUSIONS AND FUTURE WORK

We look forward to a robust discussion regarding a) our approach to automatically analyzing these interwoven heterogeneous data and especially our method of aligning vital signs to verbal summaries, and b) our proposed visualization enhancements, especially our enhancement that allows a nurse to create a narrative out of temporally-evolving patient data. While our initial results indicate the effectiveness of the relatively straightforward mapping of our automatically generated clinical events to visual alerts, we need to more thoroughly evaluate whether the additional enhancements are worthwhile. Our hope is that the integration of heterogenous patient data coupled with novel visualization tools will be an effective way to transmit a narrative of patient health. Indeed, whether or not this narrative is perfectly accurate, it should adequately engage with and highlight the most relevant patient data so that another domain expert (i.e., the receiving nurse) can quickly validate (or invalidate) it and thereby become aware of potential patient risk. Finally, we hope that this additional level of annotation (via the creation of the narrative and its validation) itself becomes data that can be taken advantage of in future analyses that promote accurate interpretation of and communication about patient health.

REFERENCES

1. GitHub. https://github.com, Accessed August 2013.

2. Google Code. https://code.google.com, Accessed August 2013.

3. An, J., Wu, Z., Chen, H., Lu, X., and Duan, H. Level of detail navigation and visualization of electronic health records. In *Biomedical Engineering and Informatics (BMEI), 2010 3rd International Conference on*, vol. 6, IEEE (2010), 2516–2519.

4. Arguello, M., Des, J., Perez, R., Fernandez-Prieto, M., and Paniagua, H. Electronic health records (ehrs) standards and the semantic edge: A case study of visualising clinical information from ehrs. In *Computer Modelling and Simulation, 2009. UKSIM'09. 11th International Conference on*, IEEE (2009), 485–490.

5. Bertram, D., Voida, A., Greenberg, S., and Walker, R. Communication, collaboration, and bugs: the social nature of issue tracking in small, collocated teams. In *Proceedings of the 2010 ACM conference on Computer supported cooperative work*, ACM (2010), 291–300.

6. Carrington, J. M. Development of a conceptual framework to guide a program of research exploring nurse-to-nurse communication. *CIN: Computers, Informatics, and Nursing 30*, 6 (2012), 293–299.

7. Carrington, J. M. The usefulness of nursing languages to communicate a clinical event. *CIN: Computers, Informatics, and Nursing 30*, 2 (2012), 82–88.

8. Carrington, J. M., and Tiase, V. L. Nursing informatics year in review. *Nursing administration quarterly 37*, 2 (2013), 136–143.

9. Cebul, R. D., Love, T. E., Jain, A. K., and Hebert, C. J. Electronic health records and quality of diabetes care. *New England Journal of Medicine 365*, 9 (2011), 825–833.

10. Chuang, J., Ramage, D., Manning, C., and Heer, J. Interpretation and trust: designing model-driven visualizations for text analysis. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, ACM (2012), 443–452.

11. Cramer, H., Evers, V., Ramlal, S., Van Someren, M., Rutledge, L., Stash, N., Aroyo, L., and Wielinga, B. The effects of transparency on trust in and acceptance of a content-based art recommender. *User Modeling and User-Adapted Interaction 18*, 5 (2008), 455–496.

12. Demner-Fushman, D., Chapman, W. W., and McDonald, C. J. What can natural language processing do for clinical decision support? *Journal of biomedical informatics 42*, 5 (2009), 760–772.

13. Effken, J., and Carrington, J. Communication and the electronic health record: Challenges to achieving the meaningful use standard. *Online Journal of Nursing Informatics (OJNI) 15*, 2 (2011).

14. Effken, J., and Carrington, J. Strengths and limitations of the electronic health record for documenting clinical events. *CIN: Computers, Informatics, and Nursing 29*, 6 (2011), 360–367.

15. Forbes, A. G., Savage, S., and Höllerer, T. Visualizing and verifying directed social queries. In *Proceedings of IEEE Workshop on Interactive Visual Text Analytics* (Seattle, Washington, October 2012).

16. Greenhalgh, T., Stramer, K., Bratan, T., Byrne, E., Russell, J., and Potts, H. W. Adoption and non-adoption of a shared electronic summary record in england: a mixed-method case study. *BMJ 340* (2010).

17. Hyun, S., Johnson, S. B., and Bakken, S. Exploring the ability of natural language processing to extract data from nursing narratives. *Computers Informatics Nursing 27*, 4 (2009), 215–223.

18. Kohn, L. T., Corrigan, J. M., Donaldson, M. S., et al. *To err is human: building a safer health system*, vol. 627. National Academies Press, 2000.

19. Munzner, T. A nested model for visualization design and validation. *IEEE Transactions on Visualization and Computer Graphics (TVCG) 15*, 6 (November-December 2009), 921–928.

20. Needleman, J., Buerhaus, P., Mattke, S., Stewart, M., and Zelevinsky, K. Nurse-staffing levels and the quality of care in hospitals. *New England Journal of Medicine 346*, 22 (2002), 1715–1722.

21. Och, F. J., and Ney, H. A systematic comparison of various statistical alignment models. *Computational linguistics 29*, 1 (2003), 19–51.

22. Plaisant, C., Mushlin, R., Snyder, A., Li, J., Heller, D., and Shneiderman, B. Lifelines: using visualization to enhance navigation and analysis of patient records. In *Proceedings of the AMIA Symposium*, American Medical Informatics Association (1998), 76.

23. Pu, P., and Chen, L. Trust-inspiring explanation interfaces for recommender systems. *Knowledge-Based Systems 20*, 6 (2007), 542–556.

24. Rind, A., Wang, T. D., Aigner, W., Miksch, S., Wongsuphasawat, K., Plaisant, C., and Shneiderman, B. Interactive information visualization to explore and query electronic health records. *Foundations and Trends in Human–Computer Interaction 5*, 3 (2013).

25. Shneiderman, B. Creativity support tools: Accelerating discovery and innovation. *Communications of the ACM 50*, 12 (2007), 32.

26. Sinha, R., and Swearingen, K. The role of transparency in recommender systems. In *CHI'02 extended abstracts on Human factors in computing systems*, ACM (2002), 830–831.

27. Stevenson, J. E., Nilsson, G. C., Petersson, G. I., and Johansson, P. E. Nurses experience of using electronic patient records in everyday practice in acute/inpatient ward settings: A literature review. *Health Informatics Journal 16*, 1 (2010), 63–72.

28. Surdeanu, M., Ciaramita, M., and Zaragoza, H. Learning to rank answers to non-factoid questions from web collections. *Computational Linguistics 37*, 2 (2011).

29. Wang, T. D., Plaisant, C., Quinn, A. J., Stanchak, R., Murphy, S., and Shneiderman, B. Aligning temporal data by sentinel events: discovering patterns in electronic health records. In *Proceedings of the SIGCHI conference on Human factors in computing systems*, ACM (2008), 457–466.

30. Wang, T. D., Wongsuphasawat, K., Plaisant, C., and Shneiderman, B. Visual information seeking in multiple electronic health records: design recommendations and a process model. In *Proceedings of the 1st ACM International Health Informatics Symposium*, ACM (2010), 46–55.