What "To-Do" with Physician Task Lists: Clinical Task Model Development and Electronic Health Record Design Implications

Daniel M. Stein, MD, MA[1], Jesse O. Wrenn, MPhil[1], Peter D. Stetson, MD, MA[1,2], Suzanne Bakken, RN, DNSc[1,3]
[1]Department of Biomedical Informatics, [2]Department of Medicine, [3]School of Nursing, Columbia University, New York, NY

Abstract

Clinical task, or "to-do" lists are a common element in the physician document known as signout. Such lists are used to capture and track patient care plan items, supporting daily workflow and collaborative patient management continuity across care transitions. While physician task lists have been shown to be important to patient safety, the tasks themselves have not been systematically examined for their subject matter, structure, or components. A manual sublanguage analysis of 500 signout tasks was conducted, and a hierarchical conceptual model for clinical tasks was inductively constructed. Tasks were classified by action type (Assess, Order, Communicate, Perform) and corresponding components. The most common task action types were Assess and Order. The most common task components were "What" type components such as Tests, including subtypes Laboratory and Imaging. This study yielded several important design considerations for future electronic health record systems that support collaborative clinical task management.

Introduction

Electronic health records (EHRs) have been touted as a way to improve informational continuity and ultimately patient care and safety by supporting health information exchange[1]. EHR support of management continuity, that is, to facilitate collaborative care by coordinating care plans generated by a variety of clinicians, is less mature. Patient care, especially in the hospital, is increasingly collaborative and frequent interruptions and shift-changes disrupt communication in this setting[2-4]. This poses a proven risk to care quality and safety, and suggests the need to develop novel tools to support safe, effective coordinated care[5]. To develop an EHR-based collaborative task manager therefore seems a fitting and timely challenge for clinical informaticians. This study aims to take a first step towards this goal: to study clinical task, or "to-do" items and construct a conceptual model that can be used to inform the design of such a system.

Background

Collaborative plans for inpatient care are typically generated through interpersonal discussions among teams of clinicians and consultants caring for an individual patient. These discussions are often distilled to lists of discrete care plan tasks, or "to-do" items written on an informal clinical document known as signout. Signout is used by physicians in the hospital to track important clinical information on each patient including a summary of the past history and present illness, a medication list, and very often a task list. The document is used during the day by the primary care team members as a reference and scratchpad for daily workflow support, and it is used at the end of shift and throughout periods of cross coverage to support patient care transfers[6]. A recent study by Horowitz et al. stresses the importance of task lists to patient care and safety[7]. The investigators showed that signout "inadequacies," such as missing or poorly formed tasks, had a detrimental effect on patient care. This confirms the clinical importance of task lists, and stresses the need for a systematic examination of signout tasks to determine how the current generation of clinical information systems could support collaborative task management. Currently, signout documents and the task lists they contain are not typically integrated into EHRs[8]. If they are digitized, they are often stored as unstructured texts and in stand-alone systems, preventing the re-use of the data and integration with alert systems or decision support. This represents a missed opportunity to support management continuity with clinical informatics methods.

We have previously described the collaborative use of an electronic physician signout document that is part of the web-based clinical information system at our medical center[9]. A frequent component of the signouts we have studied is a "to-do" list at the end of each note. While task lists are often hand-written on paper, a sample of electronic signouts from our inpatient medicine service demonstrated abundant digital representation of tasks, which we extracted and used for the study.

Methods

Study Design and Primary Data Collection: This is a retrospective, descriptive study of clinical tasks as written in signout documents. The patient population was that of patients admitted to the teaching internal medicine service of NewYork-Presbyterian Hospital, an urban academic quaternary care medical center. We collected clinical task items generated during 200 patient admissions of three or more days duration. Primary data collection was through a query of our clinical data warehouse that compiled all electronic notes written for each admission.

Data Processing and Task Extraction: The entire collection of clinical notes from the 200 admissions was analyzed to determine which contained electronic signout notes. A script written in the Ruby programming language was used to find these signout notes, and extract discrete task items as represented by lines of text that begin with a checkbox made from brackets (e.g. [ ] Check hematocrit at 9PM). Two sets of these tasks were created for use in this study. The first was a "training set" of tasks used to develop and refine a conceptual model to represent tasks. The second was a "test set" of tasks used to validate the model and to quantitatively characterize the sampled clinical tasks.

Task Model Development with the Training Set: Development of the task model was a manual, iterative process of sublanguage analysis conducted on the training set. Sublanguage analysis involves characterizing the semantic classes of the words used in a specialized domain, and analyzing the relationships or groupings of the classes as they occur in a sample corpus of the sublanguage[10]. The clinical tasks in the training set were broken into conceptual units through an inductive process, and a hierarchical list of all observed units was compiled. For example, the task: "[ ] Check hematocrit at 9pm" would be broken into the concepts: [Assess > Once], [Test > Laboratory] and [Time > At]. Some of the sample was determined to be "non-tasks," i.e. the script properly extracted them because they were lines of text preceded by a checkbox, but there was nothing in the text that indicated anything to be done. Furthermore, there were some tasks that were ambiguous and could not be classified without other contextual information. These non- and ambiguous tasks were not used in the model formulation. The model development process was iterative, as each new task examined had the possibility of affecting the hierarchy of concepts in a way that would require the previously reviewed tasks to be re-examined. Formal definitions of each concept were written, and checked by SB and JW for logical consistency.

Task Model Validation with the Test Set: Using the nodes of the hierarchical conceptual model as "tags," each task in the test set was manually tagged with matching concepts from the hierarchy. Non- and ambiguous tasks were again excluded from the analysis. If a new concept that was not represented in the hierarchy was encountered, it was noted and added to the hierarchy. All ancestors of a given tag in the hierarchy were automatically tagged – for example if the tag "Call" was applied to a given task, that task would also automatically receive the tags "Communicate" and "Synchronous" (tree structure: [Communicate > Synchronous > Call]). Every task was tagged with one and only one major action type. Tasks that were "multiples," meaning there was more than one actionable task in the extracted line of text, were split and analyzed individually. The tagged data were then processed using another script to generate descriptive statistics regarding the frequency of the task concept types and the proportions of sub-types in relation the conceptual model.

Results

Task Extraction: The task extraction script identified 122/200 admissions with at least one signout note that contained an extractable task item. The script yielded a total of 2,196 tasks. This is an average of 18 unique tasks extracted per admission. A training set of 200 tasks and a test set of 300 tasks were randomly sampled for analysis.

Task Model Development with the Training Set: After 4-5 cycles of iteration and three major and several minor structural changes to the model, the concept hierarchy could consistently and logically classify all conceptual units from each task in the training set. A clear pattern emerged in the model that indicated that the tasks could be classified from two perspectives. First, all tasks represented an action that someone needed "to-do." Therefore all tasks had some type of action by which they could be classified. The four major types of the action branch of the hierarchy were: "Assess" (actions that are performed primarily through cognitive processing of clinical data); "Order" (actions in which clinicians place an order for someone else to do, e.g. a particular treatment, test or protocol); "Communicate" (actions involving the interpersonal exchange of information), and "Perform" (actions that are carried out personally to completion by the clinician, e.g. a bedside procedure or discharging a patient from the hospital). These action types were broken down into subtypes when appropriate, for example, "Communicate" actions could be synchronous (e.g. calling someone) or asynchronous (e.g. paging someone or documenting).

For each task of a given action type, its constituent components could be described. Task components were classified as members of the types: "Who," "What," "Where," "When," "Why" and "How." These major component classifications were broken into subtypes where appropriate as well. For example, "What" subsumed the subtypes, "Documentation," "Drug," "Organization," "Procedure," and "Test." Each of these sub-sub types was also decomposed when appropriate. A representation of the conceptual model is shown in Figure 1, with a magnified view of the major classification branches (right), represented as the dark grey nodes in the complete, unlabeled hierarchy (left). Unlabeled nodes are actual concepts extracted during model development, but cannot be shown at this scale.

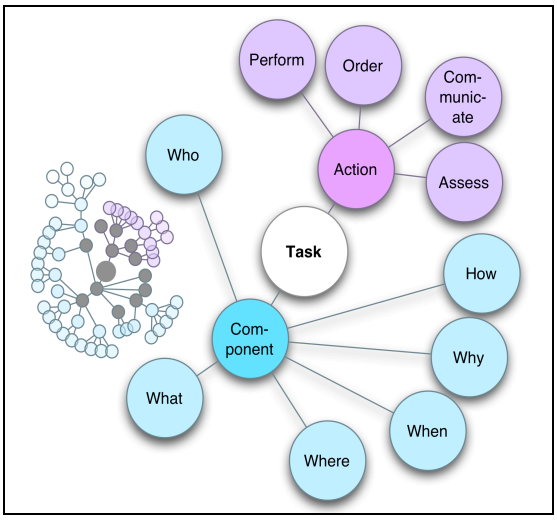

Figure 1. Hierarchical Task Model

Task Model Validation with the Test Set: 40/300 of the original test set tasks were determined to be non-tasks, and 9/300 tasks were too ambiguous to classify. There were 34 "multiple" tasks that were manually split. Subtracting the non- and ambiguous tasks and adding the extras from the split multiples resulted in a total of 286 tasks for tagging instead of the original 300 that were extracted for the test set. A total of 1,831 tags were ultimately applied to these tasks and analyzed.

No major structural changes in the model generated from the training set were necessary when tagging the test set tasks. The only changes necessary were the addition of three terminal, or "leaf" nodes to the hierarchy because a new subclass of a task component was encountered. These nodes were: [What > Test > Pathology], [What > Documentation > Flowsheets], and [What > Documentation > Discharge Summary].

Tags were processed to generate Figure 2, which shows the absolute tag counts of the: (A) action types (B) major task components, and (C) top task sub-components for all tagged tasks.

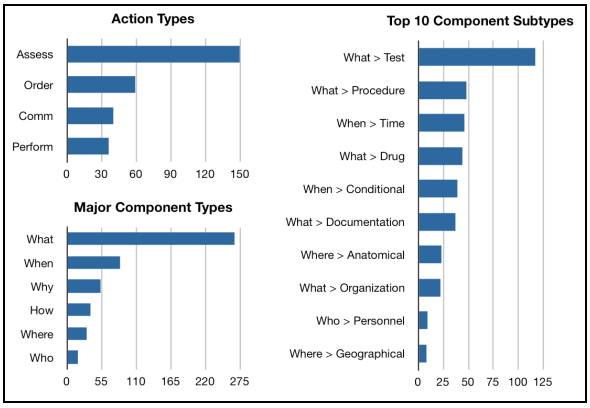

Figure 2. Absolute Counts of Actions & Components

In terms of task action type, "Assess" was the most common, more than twice as common as the next most common action type "Order." "Communicate" and "Perform" actions were the least common. "What" was the most common task component. "When," "Why," "How," "Where" and "Who" components followed in order of decreasing frequency. The sub-component [What > Test] was by far the most common, with sub-types "Laboratory," "Imaging," "Physical Measure," and "Electrophysiology" in order of decreasing frequency.

Calculating frequency by action type (as opposed to all tasks) reveals the most common task components for "Assess," "Order," "Communicate," and "Perform" tasks (Table 1).

Action Type	Task Components (% by Action Type)					
	Who	What	Where	When	Why	How
Assess	0%	100%	15%	13%	10%	5%
Order	0%	100%	2%	70%	9%	34%
Comm	43%	53%	5%	23%	73%	5%
Perform	0%	100%	17%	42%	11%	19%

Table 1. Task components by Action Type

Of the task components, all Assess, Order, and Perform tasks had at least one "What" component. Assess tasks rarely had much more than a "What" component, whereas "Order" tasks frequently had "What and "When" components. "Communicate" tasks often had a "Why" component (indicating topic for discussion) and they were split between "Who" and "What" in terms of communication targets. The

most common communication targets were [What > Organization] (which was composed of various hospital consult services) and [Who > Personnel] (which included subtypes "Physician," "Nurse" and "Social Worker").

Discussion

Creation of the model itself demonstrated a useful distinction between task action types and task components. Tasks could always be classified as one of four major types of action (Assess, Order, Communicate, Perform). In fact, it was clear that tasks without an action were not actually tasks. The "five W's and one H" approach (Who, What, Where, When, Why, How) was a useful way to categorize what would otherwise be a jumbled list of task components. Analysis of the test set tags revealed the high frequency action types, task components, and combinations of action/components. These findings provide an interesting cross sectional view of the nature of inpatient physician workflow, at least as represented in our institution's signout notes. The most common action types were "Assess" and "Order," and their respective "What" components were most frequently tests such as laboratory and imaging studies. This suggests that a major part of physician workflow on our Medicine service involves solitary, thoughtful review of patient data and placing orders, as opposed to communicating with staff or patients, or performing procedures. Beyond insight into local workflow, the findings suggest that if an EHR-integrated task manager is to be developed, it should have strong tie-ins to other aspects of the system such as laboratory and radiology results and computerized physician order entry.

Several of the other common concepts suggest design strategies for an EHR task manager. For example, the high prevalence of "When" components and more specifically [When > Time] subcomponents suggests that many tasks are time-dependent. Unlike the non-interactive format of commonly used handwritten or printed task lists, a task manager could provide active reminders or alarms for tasks that are nearly due or overdue. Another interesting frequently observed set of tags was the combination of an "Assess" type action with the "What" component "Document." Assessing a document was a frequent occurrence (36/150 Assess type actions in the test set), usually in the format: "Check for [various consult service or private attending] note." This finding suggests that much time is spend checking, and re-checking for the existence of a new note in the medical record. This type of inefficient activity could easily be avoided by the implementation of an RSS-feed-like technology that notifies clinicians of relevant additions to a patient's chart. This technique is widely used on the Web to reduce the need to frequently check sites for new updates, and allows for aggregated lists of updates of interest.

Associations between task action types and specific task components as shown in Table 1 suggest that a task manager's entry system could have intelligent entry templates that query the user for common components of tasks associated with a particular action type. For example, when entering an order task, the system may request that the user enter the commonly associated "When" and "Why." By integrating these findings into such entry templates, we could directly address the finding by Horowitz et al. that signout tasks are often missing important contextual information and putting patient care at risk.

Manual separation of the "multiple" task items allowed for an important discovery. Because of the random sampling design of the study, we generally did not see tasks in relation to their neighbors. But when multiple tasks were clumped together behind a single checkbox, we noted that while some of these groups of tasks were unrelated, several were strongly related and were, in fact, written in order and dependent on each other for their completion. For example, [Assess > Test > Lab] tasks were often followed by [Order > Drug] + [When > Conditional] tasks – e.g. check INR, dose warfarin 5mg x1 for INR <2.5. These "stacked" tasks suggest that a task manager must be able to link task dependencies. It may even be useful to obscure or hide tasks are pending other tasks' completion.

A final, revealing finding was from a casual inspection of the "non-task" items. A review of the content of the non-tasks reveals that most were notes or reminders (such as "Patient has DNR order," or "Consider starting patient on drug XYZ"). This begs the question, why is this the place physicians are putting these reminder items? A possible answer is that the signout task list is such a regular and active part of daily workflow that it serves as an excellent memory aid – perhaps it's the clinical equivalent of tying a string on your finger. Taken a step further this suggests that, given their centrality to workflow and attention, clinical task lists would serve as an excellent venue for clinical alerts and automated decision support.

One limitation of the study due to the automated task extraction method is that there are two types of tasks generally written on signouts. One type is written by and for the primary care team members to help organize tasks to be completed during the workday. The other type is written specifically for patient

handoff to help keep cross-covering physicians on top of the things they need to get done for each patient. It was not possible to distinguish these two types as tasks were extracted without context and from notes written at all times of the day.

Another limitation is that we exclusively studied physician tasks despite the fact that there are many other types of clinicians using task lists to help get things done. Specifically, nurses use task lists extensively and we plan to compare and contrast physician and nurse task types. Previous work in nursing research suggests substantial similarities. For example, the International Standards Organization (ISO) standard for a reference terminology model for nursing actions includes the concepts: "Action," "Target," "Site," "Means," and "Recipient of Care"[11]. Moreover, several nursing intervention classifications organize actions into four major action types: "Assess," "Perform," "Teach," and "Manage"[12-14]. There is only one difference between these and our action types, with "Teach" substituting for "Order." This is not surprising given the differences between practices of physicians and nurses.

Planned future work on our task model includes tagging of a larger test set by multiple domain experts, further formal validation of the model, and a comparison between physician tasks and nursing tasks.

Conclusion

We created a hierarchical, conceptual model of tasks found in physician signout notes that was well suited to classify a test set of these tasks. The exercise yielded quantitative measures of task type and component frequency and insight into the nature of clinical tasks and task management/handoff. It also provided hints at how to design an EHR-integrated clinical task manager that would enhance, but not disrupt, current clinical workflow.

Acknowledgements

Support by training grant T15-LM07079 (DMS, JOW), K22-LM008805 (PDS) and R01-NR008903 (SB). We thank Stephen Johnson for his input regarding sublanguage analysis methodology.

References

1. Kaelber DC, Bates DW. Health information exchange and patient safety. J Biomed Inform. 2007 Dec;40(6 Suppl):S40-5.
2. Coiera E. When conversation is better than computation. J Am Med Inform Assoc. 2000 May-Jun;7(3):277-86.
3. Hughes RG, Clancy CM. Improving the complex nature of care transitions. J Nurs Care Qual. 2007 Oct-Dec;22(4):289-92.
4. Yeager S. Interdisciplinary collaboration: the heart and soul of health care. Crit Care Nurs Clin North Am. 2005 Jun;17(2):143-8, x.
5. Bodenheimer T. Coordinating care--a perilous journey through the health care system. N Engl J Med. 2008 Mar 6;358(10):1064-71.
6. Van Eaton EG, Horvath KD, Lober WB, Pellegrini CA. Organizing the transfer of patient care information: the development of a computerized resident sign-out system. Surgery. 2004 Jul;136(1):5-13.
7. Horwitz LI, Moin T, Krumholz HM, Wang L, Bradley EH. Consequences of Inadequate Sign-out for Patient Care. Arch Intern Med. 2008 Sep 8;168(16):1755-60.
8. Horwitz LI, Krumholz HM, Green ML, Huot SJ. Transfers of patient care between house staff on internal medicine wards: a national survey. Arch Intern Med. 2006 Jun 12;166(11):1173-7.
9. Stein DM, Wrenn JO, Johnson SB, Stetson PD. Signout: a collaborative document with implications for the future of clinical information systems. AMIA Annu Symp Proc. 2007:696-700.
10. Hirschman L. Discovering Sublanguage Structures. In: Grishman R, Kittredge R, editors. Analyzing Language in Restricted Domains. London: Lawrence Erlbaum Associates; 1986. p. 211-324.
11. Bakken S, Hyun S, Friedman C, Johnson SB. ISO reference terminology models for nursing: applicability for natural language processing of nursing narratives. Int J Med Inform. 2005 Aug;74(7-8):615-22.
12. International Council of Nurses. International Classification for Nursing Practice. Geneva, Switzerland: International Council of Nurses; 2005 [updated 2005; cited 2005 February 9]; Available from: http://www.icn.ch/icnp.htm.
13. Martin KS. The Omaha System - A Key to Practice, Documentation, and Information Management. 2nd ed.: Elsevier; 2004.
14. Saba V. The Clinical Care Classification (CCC) System Manual. New York: Springer; 2007.