



Article scientifique

Article

2014

Published version

Open Access

This is the published version of the publication, made available in accordance with the publisher's policy.

---

Knowledge crystallization and clinical priorities: evaluating how physicians collect and synthesize patient-related data

---

Pollack, Ari H; Tweedy, Carolyn G; Blondon, Katherine; Pratt, Wanda

**How to cite**

POLLACK, Ari H et al. Knowledge crystallization and clinical priorities: evaluating how physicians collect and synthesize patient-related data. In: AMIA Annual Symposium proceedings, 2014, vol. 2014, p. 1874–1883.

This publication URL: <https://archive-ouverte.unige.ch/unige:158868>

© The author(s). This work is licensed under a Other Open Access license

<https://www.unige.ch/biblio/aou/fr/guide/info/references/licences/>

# Knowledge Crystallization and Clinical Priorities: Evaluating How Physicians Collect and Synthesize Patient-Related Data

Ari H Pollack, MD<sup>1,2</sup>, Carolyn G. Tweedy, BS<sup>2</sup>, Katherine Blondon<sup>3</sup>, MD, PhD,  
Wanda Pratt<sup>2</sup>, PhD.

<sup>1</sup>Seattle Children's Hospital, Seattle, WA; <sup>2</sup>University of Washington, Seattle, WA;

<sup>3</sup>University Hospitals of Geneva, Geneva, Switzerland.

## Abstract

*Information seeking and synthesis are time consuming processes for physicians. Although systems have the potential to simplify these tasks, future improvements must be based on an understanding of how physicians perform these tasks during clinical prioritization. We enrolled 23 physicians in semi-structured focus groups discussing simulated inpatient populations. Participants documented and discussed their data gathering and prioritization processes. Transcripts were coded to identify themes and generalized process flows. Results indicate that data are collected to categorize and prioritize patients according to expected clinical course. When data do not support these expectations, or when categorization indicates potential for morbidity, physicians increase efforts to act or re-categorize patients. Unexpected clinical changes have a significant impact on the decision-making and prioritization by clinicians. A modified version of the Knowledge Crystallization Framework helps to frame this work laying a foundation to advance information displays and facilitate information processing by physicians in clinical care environments.*

## Introduction

As increasing amounts of data become available within the medical record, Electronic Medical Record (EMR) systems need to prioritize and visualize that data to help clinicians process it in their time-constrained schedules. Clinicians perform many different tasks throughout the day—some are urgent and require immediate attention, while others can be postponed. The task of identifying priorities occurs continuously, and clinicians constantly re-evaluate their list of activities as new information surfaces. Ensuring that clinicians have all the data they need to make these decisions is a significant challenge. Today, clinicians must wade through volumes of information to find meaning, assign value, and take action. It has been estimated that physicians use approximately two million pieces of data when caring for patients.<sup>1</sup> This data includes both the acquired scholastic knowledge as well individual data for each patient. Some have said that “much of the art of medicine lies in gathering this information.”<sup>2</sup> This practice needs to move away from artistry and self-discovery to a more rigorous process that limits the potential for error. The artistry within medicine should present itself in the interpretation and application of data, not the process of finding it.

It seems clear that access to additional data can lead to more informed prioritization and improved clinical outcomes; however, the costs associated with gathering and processing these data impact both clinicians and patients. The manual processes of searching for information and analyzing it to understand trends or see patterns involves too much time, especially as patients and their problems increase in complexity. Ideally, all the relevant data required to provide clinical care should be available at the time of need. In reality, the most critical data might not be readily accessible. Therefore, a clinician must use their judgment to decide when they have enough data to proceed with analysis and decision-making, as the cost of continual searching becomes too great.<sup>3,4</sup>

In this exploratory study, we examined how physicians collect, process, and utilize data during the clinical prioritization process. We sought to model this process and examine how our model compares to the knowledge crystallization framework previously described by Card et al., which was proposed as a means for understanding and amplifying cognition during information-seeking activities.<sup>5</sup> Our goal was to identify opportunities to improve the delivery and utilization of information in the clinical care environment and future medical information systems.

## Methods

### *Recruitment & Focus Group Structure*

This study was reviewed and approved by the Institutional Review Board at Seattle Children's Hospital (SCH). Our target subject enrollment was 16 to 24 physicians. In order to be included in this study, we required that physicians be credentialed at SCH at either the attending (supervising physician) or fellow (sub-specialty trainee) level, and that they provided acute clinical services to hospitalized patients within the four months leading up to their scheduled

focus group session. Resident physicians commonly work in shifts at SCH and their prioritization process is heavily influenced by the handoff from the previous resident on duty, hence they were excluded from our recruitment efforts. We also excluded surgical providers given differences in workflows between the medical and surgical specialties.

Recruitment letters were sent to individual medical staff members. Respondents were scheduled to participate in one of five focus groups, each with 4 to 5 participants. Each subject participated in only one focus group, and we ensured that each focus group had a diverse group of participants, not limited to a single medical specialty.

All participant interactions, with the exception of recruitment, occurred during the focus groups, which took place between September and December 2013. To ensure balanced participation by all participants, each focus group was led by a focus group facilitator from our research team. Each session was 90 minutes in length and leveraged a semi-structured format divided into two parts.

During the first half of each session, participants received a series of six fictional cases (Table 1) which represented common clinical presentations and were designed to simulate a typical inpatient population. Cases were written in purposefully vague language to allow for the potential of a large differential diagnosis. Using the cases to stimulate thought, the participants wrote responses to a series of questions (Table 2) on individual worksheets. Participants were asked to provide basic demographic information on this worksheet, including age, gender, professional level (attending or fellow), medical specialty, and year of completion in medical school and graduate medical education.

**Table 1. Focus group cases with admitting or presentation diagnosis**

1. **Asthma:** Allison is 3 year old female with a history of 2 days of upper respiratory symptoms, and 1 day of increased work of breathing. At presentation to the emergency department had a significantly elevated respiratory score, and partial response to inhaled bronchodilators.
2. **Hypertension and headaches:** Ben is a 15 year old previously healthy male with 1 month of progressively worsening headaches and general malaise, who is found to have a blood pressure of 192/120 mmHg.
3. **Cellulitis:** Alex is a 6 year old female previously healthy who sustained an insect bite 3 days ago on her lower leg. She presents with 1 day of increasing redness, swelling, and tenderness of her calf and now has streaking extending up to her thigh and a temperature up to 39.5 C.
4. **Gastroenteritis:** Sara is a 3 year old previously healthy female with 4 days of vomiting and diarrhea, afebrile, unable to tolerate any oral intake, and no appreciable urine output over the past 12 hours and lethargy.
5. **Failure to thrive:** Isaac is a 4 week old infant male likely full term (exacts dates unknown, due to little prenatal care) weighing 3.28 kg and at admission weighs only 3.3 kg.
6. **Abdominal pain:** Matt is an 11 year old previously healthy male with acute onset abdominal pain, nausea, and fever.

**Table 2. Focus group worksheet writing prompts**

1. Describe the process you would use to prioritize this group of patients. In your description please consider the data that is important in the process, where this data comes from, and how you acquire this data. Please do not rank the patients in these cases.
2. Please describe how data in different situations may lead to different prioritizations. For example a fever in a previously healthy 6 year old with upper respiratory symptoms, is different from a fever in a 6 year old on hemodialysis with a central line, which is different from a fever in 6 year old with acute lymphoblastic leukemia and neutropenia. Think about how you contextualize and bundle data elements to help with the prioritization process.
3. Please list any challenges you might experience during the process you described above.

In the second part of the focus group, individual participants presented and discussed their written responses with the focus group. We recorded each focus group session with a digital audio recorder. The audio files and worksheets were transcribed for later processing and analysis.

*Data Analysis*

All participant data was de-identified prior to analysis. We coded focus group transcripts to identify themes and map out individual process flows described in the focus groups. Coding was done iteratively, and focused on participants' description of their prioritization processes. We made reference to the worksheets to verify patterns and process flows, and sought to explain differences in said flows using the demographic information provided by participants.

## Results

### Participants

We recruited 23 physicians—19 attending physicians and four fellows—representing six different medical specialties (Table 3). The participants from subspecialty services provide care to patients on their own medical service and to other services in a consultative fashion, with the exception of five participants who were Infectious Disease physicians that provide only consultative professional services.

After mapping out the process described by each participant, it became clear that physicians consistently perform the same activities during their prioritization processes. Two distinct, though similar, prioritization workflows emerged from this analysis (Figure 1).

### Workflows

The two workflows we identified contain similar elements: categorizing each patient, gathering data, and testing the fit of the collected data with the working diagnosis. The distinction between the workflows seems to correlate with a physician's identification as either one with primary patient responsibilities or one with only consultative services. Both groups employ a tacit understanding of their patients and categorize patients based on this understanding. However, providers with primary responsibility for their patients prioritize their patients based on level of concern, which determines the order of data collection for these patients. Providers collect data, test the fit of that data within their existing mental model, and then reprioritize their work based on this test of fit. It appears as though consulting providers (without primary patient responsibilities) do not prioritize their patients based on their initial clinical categorization, but gather data first, then prioritize patients based upon the fit of data within their working diagnosis.

Here we describe the two workflows using scenarios that capture the essence of participants' processes, followed by a series of representative quotes from the focus groups.

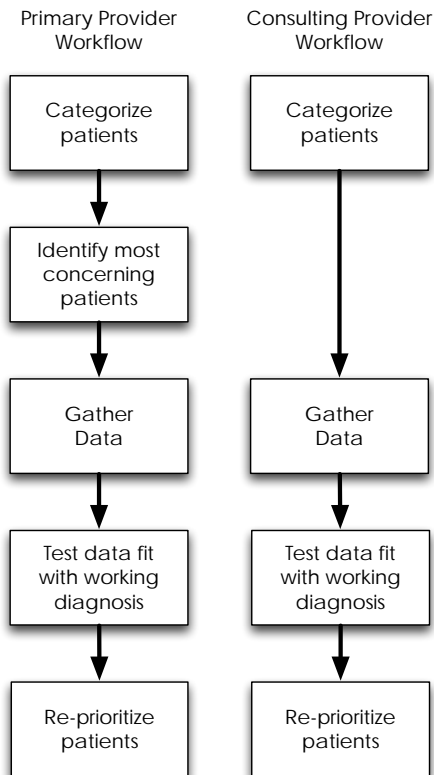
#### Workflow 1: Primary providers

*When I come into the hospital in the morning, I begin with the patients on my service for which I have the most concern. Typically these are the sickest patients, but not always. Sometimes I am most worried about the patients who I just haven't figured out or who are not behaving as expected, based on the information I have at the time I start my day. This data comes from a variety of sources, including communications received from other providers on the care team as well as the medical record. However, at this point I am mostly prioritizing my patients based on my own tacit understanding and categorization.*

*Once I have identified the patients I am most concerned about, I begin collecting additional data to investigate how that patient has changed since I last saw them. I am less concerned with absolute values (for example, an elevated creatinine level) since many were abnormal in the first place. The new data either confirms my working diagnosis or identifies inconsistencies that need further exploration and explanation. I am reassured by patients who are progressing as expected, even if they are clinically sicker, but receiving appropriate therapy. When I identify a patient who doesn't seem to fit within my mental model or categorization, I become more concerned and this patient escalates on my prioritization list. This leads me to gather additional data for this patient in an attempt to identify the correct diagnosis and ultimately provide appropriately-targeted therapy. Their response to therapy provides additional data that either supports or conflicts with my understanding of their disease process and accordingly affects my level of concern and ultimately my prioritization. Once I believe my patient to be receiving the correct therapy, I usually move on to the next patient.*

**Table 3. Provider Characteristics**

|   | Fellow  | Attending   |
|---|---|-------------|
| Number  | 4   | 19          |
| Female n (%)  | 1 (25)  | 8 (42)      |
| Age Range   | 30 -39  | 30 -59      |
| Average number of years since completing medical school (range)             | 4.5 (4-5)   | 18.1 (7-31) |
| Average number of years since completing graduate medical education (range) | NA  | 10.3 (0-26) |
| Medical specialties (n)   | General Pediatrics (4)<br>Gastroenterology (3)<br>Infectious Disease (5)<br>Nephrology (7)<br>Pulmonology (1)<br>Rheumatology (3) |             |



**Figure 1.** Workflow elements and order for primary and consultant providers.

**Supporting examples from focus groups:**

*Participant 3.2:* "...I prioritize first based on if the patient's on my service because then I'm primarily responsible as opposed to just being a consultant, and then if I'm primarily responsible, then I guess I kind of think of ... the least stable or the most sick.... [the sickest patient is] unpredictable – the patient has been primarily unpredictable while they've been there, they're not following a course we would like them to follow..."

*Participant 3.1:* "...we inherently are constantly thinking about the sickest patient more than anybody else, so when I come into the door, I automatically will open their chart up first..."

*Participant 4.2:* "...if anything catches my eye that would be worrisome, I might do a little bit more investigating."

*Participant 5.5:* "I usually want to know exactly what interventions have been tried already, because that also modifies my probability of what's going on...and kind of like just how certain we are about things, and certainly there [are] vital sign changes or [as] abnormal lab results come back, then that changes my diagnostic impression."

**Workflow 2: Representative of physicians with only consultative service responsibilities**

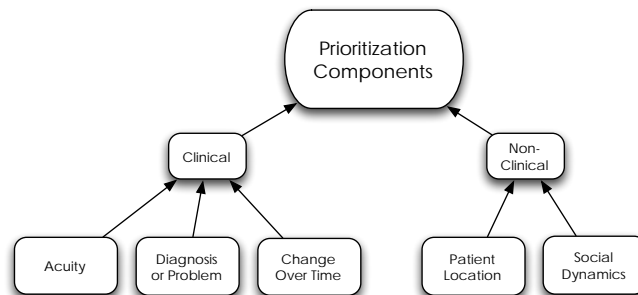
*Since I am not the primary responsible provider for the patients I am providing care to, I am less focused on their clinical acuity as I expect their primary service to address any urgent needs. Therefore, I typically begin my day with my list of patients, sorted alphabetically by last name or by location within the hospital, and then start at the top of that list and work my way down collecting new data for each patient, for example microbiology results or fever curves. As I collect these data, I am mentally processing to ensure that it fits with my understanding of the patient. I am reassured when the data fits my mental model, and more concerned when it does not. When the data doesn't fit, I spend time searching for and collecting data to help me identify the correct diagnosis. I am also concerned when I feel that the patient isn't receiving appropriate therapy, and I work with the primary team to explore why.*

**Supporting example from focus groups:**

*Participant 5.3:* "I think a couple of things that are really different about an ID consultant. One is this idea that – ah, someone else will take care of the patient. I'm totally not interested in: "are they coding at the moment"... there's other people to deal with that, so I'm expecting that if they have questions for us, they'll call us. So I'm really more concerned with more bigger picture stuff...And so trying to keep track of all of that can be one of the bigger challenges...What I tend to do when I come in...I look at everything, because what I found was that unless I looked at seven different sources of information, there were gaps."

**Prioritization: Acuity vs. Change**

When asked how they prioritize their patients, almost all focus group participants discussed invoking the concept of "sick vs. not sick" during their patient care regimes. This concept is thought to have been first used to quickly triage injured soldiers during the French Revolution, and is widely attributed to Dominique Jean Larrey, a surgeon who attended to Napoleon's Imperial Guards.<sup>6</sup> Participants in our focus groups had a difficult time defining this concept. When pushed to explain what they meant by "sick vs. not sick," participants ultimately agreed



**Figure 2.** The components used by clinicians in the prioritization process contain both clinical and non-clinical elements.

upon a process rather than a definition; a process performed by experts with expert knowledge. Physicians often use experience and intuition in conjunction with objective data to decide who is sick. Most participants cited clinical acuity as the basis for the distinction, including a variety of data points most commonly including vital signs and laboratory values. Clearly, clinical acuity plays a role in the prioritization process, but changes or trends seem to play a vital role in this process, as well (Figure 2). Changes in either objective or subjective data elements—both expected changes and unexpected changes—appear to facilitate the distinction between patients who are sick, and those who are not. Quotes from the focus group participants explain this best:

*Participant 2.2: “So I'm looking for patterns that I recognize that I'm supposed to be seeing in patients in general based on experience. I'm looking for how the patterns are not following what I'm expecting based on my knowledge of the type of patient ... And I'm looking for a change in the pattern or in the trend which maybe I wasn't expecting...And that stops me in my tracks, why is this not making sense? ... That's when everything stops and that's when I really want to be able to have the ability to grab around all the data, the numbers, the written words, the commentary which is hard to find, the imaging, the everything, and just sort of say why does this not make sense?”*

*Participant 2.4: “So if you're sure of the diagnosis and you're sure they're on the right therapy even if they're worse, not such a big deal. If all of a sudden you say, ‘Wait, this doesn't make sense anymore.’ You have to go back to square one...Are we treating the wrong thing?”*

*Participant 2.2: “... We have plenty of patients who get worse, but sometimes that's part of pattern. And that actually doesn't freak you out anywhere near as much as when somebody gets even just a little bit worse, but that's not what they're supposed to do. It's when they do what they're not supposed to do that you think something is weird going on, you have to regroup.”*

*Participant 4.1 “Change, yeah, a new vital sign change, new critical lab value. I'm trying to get at that decompensation potential...”*

Our findings suggest that physicians express elevated levels of concern and anxiety for patients that do not follow an expected clinical course. Upon encountering this scenario, participants indicate a heightened level of concern, which leads to increased vigilance for those patients who have the potential to acutely decompensate. In order to uncover this type of discrepancy, participants indicated that they rely on changes in clinical outcomes and did not find absolute values as helpful.

### Challenges

Participants expressed many concerns and difficulties, relating to both data collection and processing, when trying to prioritize their patients (Table 4 and Table 5). Participants identified missing data as a significant problem. A variety of factors contribute to missing data—it might not have been collected in the first place, or it might have been collected but stored in an alternative location (e.g. lab result in an outside record). Many participants echoed this sentiment and expressed frustration with data sources that are inaccessible or hard-to-access, especially data retained by a nurse, consultant, or patient.

In addition to missing data, another major challenge identified during the focus groups related to the time required to collect relevant data. Participants reported the need to search across multiple sources, oftentimes in a variety of physical locations. The need to search through voluminous data sources and ultimately filter out non-relevant segments significantly contributed to the amount of time participants required to complete the data collection process. Other challenges encountered related to data quality as well as technical challenges with interacting with

#### **Table 4. Challenges with Data Collection**

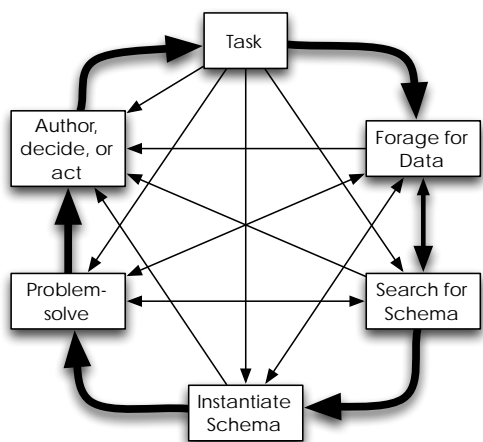
1. Missing data
  - Not collected
  - Collected but not resulted
  - Collected but not available or not found
    - Data in outside record
    - Source (i.e. resident, nurse, consultant, patient, care provider) not accessible
    - Most recent results not available
  - Available data not recognized as valuable
2. Communication between care team members did not occur
3. Noise
  - Filtering unimportant data
4. Time Consuming
  - Searching
  - Waiting (related to availability of resources)
  - Need to check multiple sources in a variety of locations
5. Quality concerns
  - Data gatherers have varying skill levels
  - Language barriers
6. EHR
  - Speed
  - Interface

**Table 5. Challenges with Data Processing**

1. Bias
  - Anchoring
  - Bandwagon
  - Confirmation
  - Premature Closure
  - Presentation
  - Source value
2. Time Consuming
3. Multitasking
4. Different focuses between stakeholders (physician, nurse, patient, caregiver)
5. Unexpected results
6. Discovering trends especially across data types
7. Framing, putting results in context
  - Normal vs abnormal
  - Better or worse
  - Over time

## Discussion

Clinicians spend much of their time working with information: searching, acquiring, processing, framing, and acting on information with the goal of providing the best patient care. Most clinicians in a hospital setting are responsible for providing care to multiple patients at once. Given the multitude of competing needs, clinicians must constantly re-prioritize their work to ensure the most urgent needs are met first. This process might best be described as a “knowledge crystallization” task, which was first described by Card et al. as a task “in which a person gathers information for some purpose, makes sense of it by constructing a representational framework (referred to as a schema), and then packages it into some form for communication or action.”<sup>3</sup>



**Figure 3.** Knowledge Crystallization framework as proposed by Card, Mackinlay, and Shneiderman in 1999. This model aids information seekers by providing structure as well as context to their search process ultimately reducing the cognitive burden needed to accomplish a task. Typically, one moves around the outside in a counterclockwise fashion, but can jump to other steps as appropriate.

the electronic medical record.

During the focus group discussions, participants also indicated that they encounter difficulties processing information after it has been collected. Many of these challenges concerned cognitive biases that interfere with the participant’s ability to appropriately process and interpret available data. For example, some participants expressed challenges with incorporating new data into existing mental models, especially when the new data did not fit or was unexpected. Despite the quantity or quality of data that may be available within the medical record, a clinician may not always think to gather all potentially relevant data, or may even miss an important test result, due to an error in their clinical reasoning or potential retrieval issues.

Other themes that emerged around the discussion of data processing were difficulties with trending data over time, across data types, and putting it in the appropriate context.

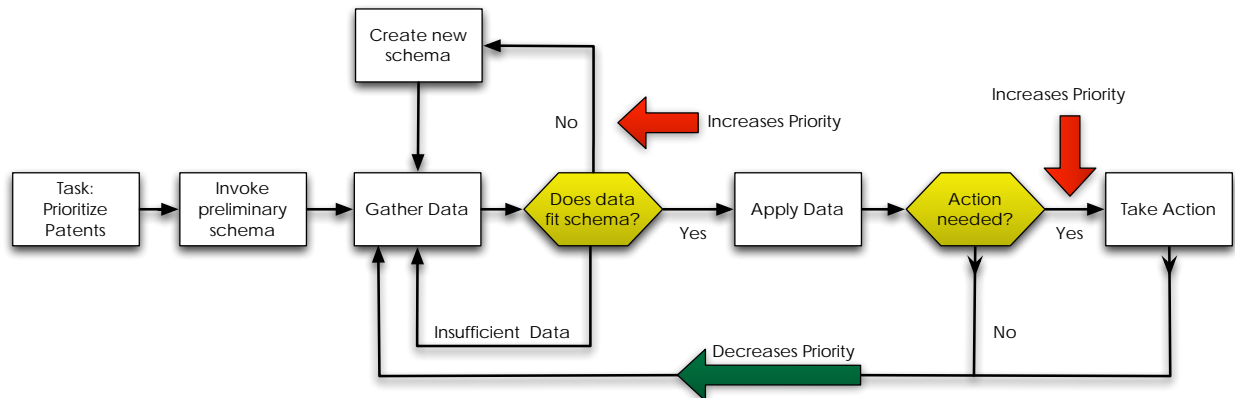
### Knowledge Crystallization

Within the knowledge crystallization framework (Figure 3), a knowledge seeker typically moves from a task to gathering data to finding and instantiating an appropriate schema that allows the individual to address the task and act appropriately. Card et al. describe this process in detail, and specify three required elements for knowledge crystallization: data, tasks, and schemas.<sup>3</sup> The value of the knowledge crystallization model lies in its ability to improve the efficiency of acquiring and processing data in order to accomplish a stated task. Initially envisioned as a tool to create schemas for information visualization models, use of knowledge crystallization extends to most information seeking activities.

At a high level, the prioritization process used by clinicians during patient care activities seems to model the knowledge crystallization framework. However, the process revealed by physicians in our focus groups indicates a slight—though notable—difference.

When caring for patients, clinicians have a primary task: provide the best and most appropriate therapy in a timely fashion. To accomplish this task, clinicians systematically collect targeted data. To be effective, clinicians begin with a mental model or schema they have deemed appropriate to their current situation, and they ask questions and seek answers that test the validity of this schema (in the field of clinical reasoning, this is analogous to the hypothetico-deductive approach, where experts use co-selection in order to test a hypothesis).<sup>7,8</sup> This process continues until the clinicians reach a point where they are comfortable enough to act and deliver appropriate care.

Therefore, we cannot accept the knowledge crystallization framework wholesale without making modifications. Recognizing and understanding the difference between the knowledge crystallization framework and a physician's prioritization process allows us to propose a new model which highlights the importance of the schema for clinicians while providing patient care. The new model resulting from this research (Figure 4) incorporates the concept of the schema and how the prioritization process is influenced by the collection and processing of data.



**Figure 4.** Modified knowledge crystallization model addresses the task of prioritizing patients. Once clinicians have invoked their preliminary schema, categorizing each patient, they gather data and assess the fit of the data with their schema. Poor-fitting data requires a new schema, increasing the priority of that patient.

Beginning with a schema prior to gathering data allows a clinician to proceed with purpose and direction, what some would refer to as the artistry in medicine.<sup>9,10</sup> When a physician determines the need to take action and provide care, therapeutic interventions appear to be based on their diagnostic categorization (i.e. schema). Depending on the acuity and urgency of the situation, clinicians may choose to act with a low level of diagnostic certainty to prevent further clinical deterioration. However, regardless of the degree of certainty with the working schema or diagnosis, instantiating a schema allows for constant evaluation and testing, generating refinements, increasing accuracy, and ultimately leads to the appropriate intervention for the patient.

Our model identifies the value of the schema and how clinicians use data to test its validity. Clinicians place high importance on the validity of their schema: when they uncover data that contradicts that schema, their level of concern for the patient, and therefore the patient's priority increases. We expected clinical acuity (i.e. sick vs. not sick) to be the major driver in the prioritization process; but this research suggests that classification errors leading to poor schema formation have equal, if not greater importance. Recognizing the importance of patient categorization and ultimately schema formation has the potential to influence future information delivery systems.

#### *Implications*

Tools designed with this new model in mind will have a greater capacity to dramatically improve a clinician's ability to provide care for their patients, and streamline their work processes. We first propose design principles to help guide the development of new clinical information system tools followed by specific recommendations where use of our model has the potential for positive change: workflow improvements to assist in patient categorization, and potential integrations with the burgeoning field of information visualization.

#### *Design Principles*

Our model provides multiple opportunities to improve clinical information systems. Designers should focus their development efforts highlighting information that contradicts the pre-formed schemas utilized by clinicians in the care process. These tools have the potential to highlight subtle and early changes in the clinical course of a patient, perhaps even before a clinician has recognized them. In addition, our research clearly demonstrated the importance and value clinicians put on expected and unexpected changes in the clinical courses their patients follow. Therefore, the development of new tools should focus on highlighting these types of changes, making it easier for clinicians to detect, process and act. Designers have the opportunity to focus their efforts on concepts such as change from expected, change over time, as well as relative changes and absolute changes.

### *Patient Categorization*

Large data repositories—a result of the implementation of electronic health records at many hospitals—have tremendous potential for machine learning applications. Machine learning involves the process of applying an algorithm to accomplish a task, where “learning” is achieved through recursively repeating the task and improving upon an empirical metric. Clustering, a typical machine learning problem, sets out to group objects such that the objects within a group or cluster have similar traits. The process of assigning a diagnosis to a patient is analogous to cluster analysis, in that a physician identifies specific traits of an individual (e.g. fever, cough, chest pain) and identifies potential similar groups or diagnoses for that patient (e.g. upper respiratory infection, pneumonia or asthma). As the physician gains more information she adjusts the grouping to improve the fit, just as we have described in our model. There are examples of using cluster analysis to better understand the variability within a specific disease, but we are not aware of its use in clinical systems to help in real-time patient classification.<sup>11</sup> The physicians we interviewed agreed on the importance of recognizing when patients begin to deviate from their expected course, and it is clear that an automated cluster analysis tool built into the electronic medical record with the potential to highlight these changes would be tremendously valuable to physicians. Furthermore, given the biases that interfere with the medical decision-making process these tools have the potential to provide more objective evidence to providers during times of great need.<sup>12</sup>

### *Information Visualization*

One could argue that the results of cluster analysis achieve their greatest potential when expressed visually. The human visual system quickly allows us to find patterns and associations within data sets when organized in a visual fashion.<sup>13</sup> Cluster analysis exploits this trait to discover patterns and groupings based on specified attributes.

Imagine a clinical system that visually represents groups or clusters of patients based on historical clinical attributes from within the medical record. Patients of an individual physician could be overlaid on this display, allowing the physician to quickly identify appropriate groupings for his or her patients. It would be possible to show how a patient’s affinity towards different groupings changes over time as new data becomes available. Our results demonstrate the importance of categorization and schema formation within clinical care, specifically how inaccurate categorization of patients leads to increased concern from medical providers. Visual presentation of cluster analysis has incredible potential to help physicians recognize subtle changes in a patient’s clinical course more quickly, reducing this anxiety.

We recognize that this vision for cluster analysis within the medical record will be challenging to implement and may be in the future but information visualization still has great potential with much simpler tools using our model as a basis for presenting information to clinicians. For example, visualizations that identify changes in patient variables over time would be easier to implement, and still offer tremendous benefit to physicians. Wang et al. indicate that tools have been developed to visualize event time sequences that provide clinicians and administrators insight into how the order of events affect overall patient care.<sup>14,15</sup> Leveraging tools to highlight changes in clinical outcomes across different data points, like changes from baseline, changes from normal, or changes from an expected course, support the task of clinical prioritization. Most electronic systems provide this basic capability via result trending over time, but it is easy to imagine how more intricate displays spanning disparate data sources could help physicians identify patients with more urgent needs. As indicated in our focus groups, physicians tend to sort their patient list by demographic information. A patient list with sort options encompassing the degree of change over the last 6 or 12 hours based on vital signs and lab values would help physicians quickly identify patients with the highest degree of change among their population of patients. Combined with a physician’s tacit knowledge of the expected course for an individual patient, this tool could help identify patients with unexpected deviations in their clinical course.

These examples leverage Shneiderman’s mantra for information visualization: overview first, zoom and filter, and then details on demand, which would provide clinicians with a quick, but highly interpretable snapshot of their patients’ statuses.<sup>16</sup> Although clinicians often create mental models to help process this information, visualization techniques could help lighten cognitive processing required by physicians as well as provide new insights into data available within the medical record.<sup>17,18</sup> Chittaro describes a series of techniques and goals for visualizing information and applying them to the medical domain.<sup>19</sup> However, none of these tools help to prioritize information within a single record, nor across multiple patient records. Leveraging information visualization to highlight the points in our model that increase a patient’s priority has the potential to dramatically improve physicians’ ability to prioritize all of their patients.

### *Strengths and Limitations*

Too often the evaluation of healthcare information technology occurs late in the software development cycle, making changes much more difficult and costly to implement.<sup>20</sup> In addition these late evaluations base their system design on assumptions of stakeholder needs and workflows. Conversely, our work explored understanding the fundamentals of information seeking and processing for practicing physicians. This knowledge should allow clinical information system designers to develop systems that better target the needs of clinicians. Because our study based its findings on a diverse sample of medical providers, the results should be generalizable across a broad population of providers.

Despite the diversity of medical providers within our sample, all of the participants were pediatric providers at a single organization, which might diminish its generalizability. Nonetheless, we would argue that pediatric providers approach the task of clinical prioritization in the same manner that internal medicine or family practitioners do. In addition, the scenarios utilized in the focus groups only included hospitalized patients, excluding the outpatient setting, though most of the participants in the study provide care in both settings. One could easily repeat this study and include an expanded group of clinicians in a variety of settings to ensure the broader applicability, especially for ambulatory patients where care extends over weeks to months and even years.

### **Conclusion**

Physicians expend a significant amount of time and mental effort working with data, finding context, and ultimately transforming information into knowledge, which proves to be a tremendously costly endeavor. Understanding the cost structure of information seeking has broad implications in healthcare information systems. Development of these systems should be based on a thorough understanding of clinical knowledge acquisition and processing as well as the costs of these procedures.

The first step in reducing these costs requires us to understand how physicians approach the journey of knowledge discovery and identify points or situations that have the potential to increase these costs. We present a model that describes the information seeking behavior used by physicians to prioritize patients in a hospital setting, highlighting the importance of accurate clinical categorization in this process. Our goal for this research was to inform the design and creation of novel tools or systems that focus on the data and information required during the prioritization process and improve their delivery. Through machine learning applications and information visualization, our model has great potential to reduce the costs of knowledge acquisition in healthcare settings and as a result improve clinical outcomes.

### **Acknowledgments**

We would like to thank all participants and funding sources (NIH T32 DK007662) and Dr. David Hendry for his support in this endeavor.

### **References**

1. Pauker SG, Gorry GA, Kassirer JP, Schwartz WB. Towards the simulation of clinical cognition. *The American Journal of Medicine*. 1976 Jun;60(7):981–96.
2. Smith R. What clinical information do doctors need? *BMJ : British Medical Journal*. BMJ Group; 1996 Oct 26;313(7064):1062.
3. Card, S. K., Pirolli, P., & Mackinlay, J. D. (1994). The cost-of-knowledge characteristic function: display evaluation for direct-walk dynamic information visualizations, 238–244. doi:10.1145/191666.191753
4. Pirolli, P., & Card, S. (1999). Information foraging. *Psychological Review*, 106(4), 643–675. 3. Card SK, Mackinlay JD, Shneiderman B. *Readings in Information Visualization: Using Vision to Think*. 1999.
5. Card SK, Mackinlay JD, Shneiderman B. *Readings in Information Visualization: Using Vision to Think*. 1999.
6. Shapiro N, Howell MD. Sick? Or, not sick?\*. *Critical Care Medicine*. 2005 May 1;33(5):1151.
7. Schmidt HG, Norman GR, Boshuizen HP. A cognitive perspective on medical expertise: theory and implication [published erratum appears in *Acad Med* 1992 Apr;67(4):287]. *Academic Medicine*. 1990 Oct 1;65(10):611.
8. Elstein AS, Schwarz A. Evidence Base Of Clinical Diagnosis: Clinical Problem Solving And Diagnostic Decision Making: Selective Review Of The Cognitive Literature. *BMJ : British Medical Journal*. BMJ; 2002 Mar 23;324(7339):729–32.

9. Schon DA. Educating the reflective practitioner: toward a new design for teaching and learning in the professions. 1st ed. San Francisco: Jossey-Bass; 1987. 1 p.
10. Montgomery K. How Doctors Think: Clinical Judgment and the Practice of Medicine. Oxford University Press; 2005. 1 p.
11. Haldar P, Pavord ID, Shaw DE, Berry MA, Thomas M, Brightling CE, et al. Cluster analysis and clinical asthma phenotypes. *Am J Respir Crit Care Med*. 2008 Aug 1;178(3):218–24.
12. Dawson NV, Arkes HR. Systematic errors in medical decision making: judgment limitations. *J Gen Intern Med*. 1987 May;2(3):183–7.
13. Tufte ER. The visual display of quantitative information. Second. Cheshire, Conn.: Graphics Press; 2001.
14. Wang TD, Plaisant C, Shneiderman B, Spring N, Roseman D, Marchand G, et al. Temporal Summaries: Supporting Temporal Categorical Searching, Aggregation and Comparison. *IEEE Trans Visual Comput Graphics*. 2009;15(6):1049–56.
15. Wang TD, Wongsuphasawat K, Plaisant C, shneiderman B. Visual information seeking in multiple electronic health records. New York, New York, USA: ACM Press; 2010. pp. 46–55.
16. Shneiderman B. The Eyes Have It: A Task by Data Type Taxonomy for Information Visualizations. *Visual Languages*. 1996;:336–43.
17. Kang Y-A, Görg C, Stasko J. How Can Visual Analytics Assist Investigative Analysis? Design Implications from an Evaluation. *IEEE Trans Visual Comput Graphics*. 2010 Jun 2;:570–83.
18. Liu Z, Stasko JT. Mental models, visual reasoning and interaction in information visualization: a top-down perspective. *IEEE Trans Visual Comput Graphics*. 2010 Nov;16(6):999–1008.
19. Chittaro L. Information visualization and its application to medicine. *Artif Intell Med*. 2001 May;22(2):81–8.
20. Yen P-Y, Bakken S. Review of health information technology usability study methodologies. *J Am Med Inform Assoc*. 2012 May;19(3):413–22.