Automated methods for the summarization of electronic health records

Rimma Pivovarov and Noémie Elhadad

ABSTRACT

Objectives This review examines work on automated summarization of electronic health record (EHR) data and in particular, individual patient record summarization. We organize the published research and highlight methodological challenges in the area of EHR summarization implementation.

Target audience The target audience for this review includes researchers, designers, and informaticians who are concerned about the problem of information overload in the clinical setting as well as both users and developers of clinical summarization systems.

Scope Automated summarization has been a long-studied subject in the fields of natural language processing and human–computer interaction, but the translation of summarization and visualization methods to the complexity of the clinical workflow is slow moving. We assess work in aggregating and visualizing patient information with a particular focus on methods for detecting and removing redundancy, describing temporality, determining salience, accounting for missing data, and taking advantage of encoded clinical knowledge. We identify and discuss open challenges critical to the implementation and use of robust EHR summarization systems.

INTRODUCTION

The increased adoption of electronic health records (EHRs) has led to an unprecedented amount of patient health information stored in electronic format. However, the availability of overwhelmingly large records has also raised concerns of information overload,[1] with potential negative consequences on clinical work, such as errors of omission,[2] delays,[3] and overall patient safety.[4] Current EHR systems often do not present this tremendous amount of patient data in a way that supports clinical workflow or cognitive reasoning.[5] It is therefore imperative for patient care to automatically comb through the raw data points present in the records and detect timely and relevant information.

Alarmingly, as the most chronically ill patients often have the largest datasets, their records are the most difficult to coherently present.[6] As an example, for a prevalent chronic condition in our institution, patients with chronic kidney disease have 338 notes on average in their record (from all clinical settings) gathered across an average of 14 years, with several patients' records containing over 4000 notes. It is clear that during a regular medical visit, no practitioner can read hundreds of clinical notes. Fortunately, electronic storage of this health information provides an opportunity for EHR systems to "aid cognition through aggregation, trending, contextual relevance, minimizing superfluous data."[7] Currently available commercial EHR systems, however, inadequately address this need, sometimes providing organization of data but lacking in information synthesis.[8] Some vendor EHR dashboards display problem lists that aggregate billing codes but these are low in actionable knowledge.[9,10]

Given this unmet and well-recognized need for comprehensive EHR summarization,[11,12] many research groups have designed and evaluated clinical data summarizers. In this review, we sample summarization applications to highlight different features including seminal work, different evaluation strategies, and various input/output data. We also examine the current work and future directions for six challenges of EHR summarization: information redundancy, temporality, missing data, salience detection, rules and heuristics, and deployment of summarization tools.

GENERAL APPROACHES TO SUMMARIZATION

There are multiple theoretical frameworks for summarization in the clinical domain[13] as well as for textual summarization in the general domain.[14,15] In the broader field of summarization, there has been a lot of work in automated text summarization, specifically within the genres of news stories and scientific articles (see[16] for an in-depth review). Clinical summarization, "the act of collecting, distilling, and synthesizing patient information for the purpose of facilitating any of a wide range of clinical tasks,"[13] presents a different set of challenges from summarization in other domains and genres of texts.

While there exist other discussions on biomedical literature summarization methods[17,18] and EHR visualizations,[19–21] in this review we focus on characterizing existing clinical summarization systems by outlining the system outputs and evaluations as well as highlighting the remaining challenges that exist in automated summarization.

To categorize the summarizers highlighted in this review, we focus on two common dimensions used in the text summarization literature: extractive/abstractive summarization, and indicative/informative summarization. We define the four categories that describe summary types.

1. *Extractive summaries* are created by borrowing phrases or sentences from the original input text. In the domain of clinical summarization, an extractive approach can identify pieces of the patient's record and display them without providing additional layers of abstraction.

Correspondence to Rimma Pivovarov, Department of Biomedical Informatics, Columbia University, 622 West 168 Street, VC-5 New York, NY 10032, USA; rp2521@cumc.columbia.edu; Tel: 781-738-5272; Fax: 212-305-3302

2. *Abstractive summaries* generate new text that synthesizes the original text. In the domain of clinical summarization, abstractive summaries may provide additional higher-level context to explain the data, such as computed quantities (e.g., trends) or automatically generated text.

Extractive and Abstractive summaries are further categorized as either indicative or informative.

3. *Indicative summaries* point to important pieces of the original text, highlighting significant parts for the reader. In the domain of clinical summarization, indicative summaries may convey, for instance, when key tests were performed or diagnoses were made. Indicative summaries are meant to be used in conjunction with the full patient record.
4. *Informative summaries* replace the original text. In the domain of clinical summarization, informative summaries are designed to be used independently of the full patient record, meaning they are used as a replacement for the original full set of raw data.

How to evaluate a summarizer, both its accuracy and its added value in supporting users carry out information-related tasks has also been the subject of investigation in general domain and clinical summarization. Intrinsic evaluations focus on the internal validity of a summarization tool. Typically, experts evaluate the quality of the automatically produced summaries; or themselves create gold-standard summaries, against which automatic ones are compared. In an extrinsic evaluation framework, the usefulness of the summarization tool is assessed through its effectiveness in helping individuals carry out a task. For instance, a clinical summary could be evaluated in an extrinsic fashion by comparing how quickly and accurately trial coordinators can identify patients eligible for a trial with access to patients' full records or with access to a summary instead.

Almost since the inception of EHRs, there has been an interest in creating meaningful succinct summaries for clinicians. The research on automated summary creation has spanned over 30 years and initiated with extracting recent structured events in a patient's history[22] evolving into performing natural language processing (NLP)[23] and automatically linking different data types[24,25] to create a more holistic view of the patient record. Table 1 lists clinical summarization systems proposed in the research literature in chronological order. We describe each system according to the following axes: the summarization approaches it implements, the type of input data it handles, the type of output summary, the way in which it was evaluated, and whether it was deployed in a clinical environment. Overall, summarization approaches investigated in clinical summarization have primarily been for indicative and extractive summarization. We also note a lack of evaluation, especially in the most recent years. We discuss in further detail the methods used for summarizing clinical data, along with the open research questions present in each of the summarization steps.

METHODOLOGICAL CHALLENGES
The following sections present some unsolved challenges in clinical summarization. A conceptual framework proposed by Feblowitz et al.[13] defines a set of actions that successful summarizers should accomplish with raw information: Aggregate, Organize, Reduce/Transform, Interpret, Synthesize. We discuss methodological challenges with automated summarization within the context of this framework.

Specifically,

– To successfully *aggregate* disparate clinical data sources, the ability to recognize and account for **similarity** is imperative. Such similarity occurs at different levels within narratives: from word-level similarity to concept to statement-level; as well as in other data types and across. We focus our discussion on textual similarity.
– The *organization* and *interpretation* of the aggregated data requires extraction and reasoning over clinical events and their **temporality**. We examine extraction of temporal information from text along with representation and reasoning over clinical events.
– The *organization* and *interpretation* of the aggregated data also requires that **missing data** points be accounted for. Patients are sometimes seen with predictable regularity but are most often seen at erratic intervals. Missing data points are often filled in by imputation, adding missing data indicators, deleting information with missing data, or other strategies.
– In the *reduction and transformation* of data and its *synthesis*, it is critical to decide which pieces of information are **important** and must be contained in the summary. Some methods for automatically detecting importance have relied on linguistic structure while others use probabilistic modeling techniques.
– To provide context for *interpretation* and *synthesis* of clinical data, it is useful to employ **existing knowledge** and create rules for the summarization. Knowledge-based heuristics often provide a way to specify time constraints, concept relationships, and abstractions.
– Finally, to successfully implement summarizers into clinical care, challenges of **deployment** need to be addressed. Because in vendor EHR systems there are limited opportunities to deploy innovative and experimental technology, there have been few attempts to translate patient record summarization systems into the clinic; however, to demonstrate utility, it is imperative to implement and study clinical summarization tools in the real world care setting.

1) IDENTIFYING AND AGGREGATING SIMILAR INFORMATION
We review approaches to identifying and aggregating similar information on three different levels of language abstraction: words, concepts, and statements, as investigated within and outside the field of clinical summarization.

Word-level Similarity
In clinical NLP, much work has been devoted to identifying lexical variants that are similar in meaning.[43] The Unified Medical Language System (UMLS),[44] for example, provides essential knowledge towards that goal by grouping words into concepts. For instance, the terms MI, myocardial infarction, and heart attack all share lexical similarity, and map to the same underlying concept. Within clinical summarization, normalization of words to concepts has only recently been investigated.[42,45]

An alternative, and most common approach in clinical summarization, is to identify word-level similarity by finding redundant strings of words. Patient records often contain redundant spans of text – this can be explained by the fact that documentation is often formulaic but also by the common habit of clinicians to copy and paste text from one note to another.[46] Multiple different automated methods have been employed to identify copy and pasted words within clinical notes. A plagiarism detection tool called CopyFind has been used to identify overlapping phrases in input texts.[47] More recently, global[48] and local[45,49] bioinformatics-inspired alignments have been proposed for

REVIEW

Table 1: A sampling of clinical summarization applications, organized by publication date

	Summarization approach	Input	Output	Evaluation	Deployed (when is it generated)	General Notes
NUCRSS[22,26]	Extraction of clinical variables, indicative	Real structured EHR data	An eight page summary of: Problem list, Vital signs, Cardiac-pulmonary-renal diagnoses, Treatments, Routine specialized laboratory examination, Suggestions to physicians regarding patient care	Laboratory study with medical students and physicians showed significant time savings and increased accuracy. Randomized controlled trial found showed that the NUCRSS improved process level (patient's length of stay and increased the amount of laboratory tests ordered) outcomes and may have improved care.	Yes (each patient visit)	Early example of a summarizer. One of the few summary evaluations that demonstrate an impact on quality of care and process outcomes.
STOR[27]	Extraction of clinical variables, indicative	Real structured and unstructured EHR data	Loosely customizable, summary which included both time- and problem- oriented views	Clinical study found that clinicians were better able to predict their patient's future symptoms and laboratory test results when the using medical record in addition to STOR as opposed to just the medical record.	Yes (each patient visit)	Early example of a summarizer. One of few examples of task-based evaluation. The summary is context-dependent on the patient, but the context is manually determined by the clinician (what problems are active, what observations are relevant, etc.).
Powsner and Tufte[1,28]	Extraction of psychiatric variables and recent notes, indicative	Simulated structured, unstructured and genealogy data	A one-page summary that visualizes the most salient content (as defined by recency) of the patient record.	None	No	A widely referenced prototype that continues to serve as a model for current EHR visualization and summarization applications.
Lifelines[29,30]	Extraction of clinical variables, indicative	Simulated structured data	Holistic interactive patient summaries using a temporal data view on top of the raw EHR data. Displays facts as lines on graphic time axis according to their temporal location and categories/significance are represented by color and thickness.	The original Lifelines application was evaluated for work with juvenile youth records[29] by a small group of users who reported enthusiasm but mentioned potential biasing by the system's graphics.	No	Lifelines is probably the most well-known summarizer tool. The display has served as a model for future timeline-view clinical summarizers. Lifelines2 was created for research and examining many patients together.
ClinViewer[23]	Extraction of concepts from text, indicative	Real unstructured EHR data	Combined NLP techniques and presented a tree view of a patient's problems extracted from the narrative text to the clinician. Displays concepts in context when clicked.	The system was able evaluated on accuracy and speed using real discharge summaries but no evaluation with clinicians was conducted.	No	One of the first examples of summaries created using NLP. Allows for customizable user views. Works on top of the MedLEE[31] NLP engine which handles modifiers.
IHC Patient Worksheet[32]	Extraction of clinical variables, indicative	Real structured EHR data	1–2 page outpatient summary of: Demographics, Problems, Medications, Laboratory tests, Actionable advisories	A retrospective cohort study found that compliance with HbA1c testing was higher for patients who had a worksheet printed than for those who did not.	Yes (each patient visit)	One of the few example of a clinical outcome tested in the evaluation
CLEF[33–35]	Abstraction from text and extraction of clinical variables, indicative	Simulated structured and unstructured cancer patient data.	An interactive display of both navigational capabilities for the EHR (indicative) and generates textual summaries (abstractive) to enhance comprehension. It uses information extraction techniques to identify classes of data and relationships between them.	None	No	One of the few natural language generation systems created for medical histories. Represents histories as a semantic network of events organized temporally and semantically. Lists requirements that are very relevant to general designers of clinical summaries – the list was generated via initial requirements elicitation process. Uses a logical model of cancer history

Table 1: Continued.

	Summarization approach	Input	Output	Evaluation	Deployed (when is it generated)	General Notes
KNAVE-II[36]	Abstraction and extraction of clinical variables, informative	Real structured data on bone marrow transplant patients	Interactive data display of abstracted and raw protocol-based care data containing a tree-browser and time chart.	A crossover study compared KNAVE-II with paper charts and Excel spreadsheet. Users produced quicker answers, had somewhat better accuracy and preferred KNAVE-II however it did not achieve a very high system usability score.	No	Performs semantic, temporal, and context abstraction. Requires domain-specific ontologies. Consists of a knowledge base, abstraction generator, navigation engine, and visualization. Lists 12 desiderata for interactive, time-oriented clinical data that should be used to guide future summarization work as well.
BabyTalk (BT-45)[37,38]	Abstraction of ICU data streams, informative	Real raw neonatal ICU data streams	Automatically generated natural language to describe ICU data streams for easier comprehension by the nursing staff.	A laboratory study found that human-generated text summaries of ICU streams helped nurses predict their patient's trajectories' better. The team is working to create automatically generated text summaries that perform as well as human-generated summaries.	No	A novel example of summarizing graphical ICU information by generating text.
Were et al.[39]	Extraction of clinical variables, indicative	Real structured EHR data from OpenMRS	Patient summary for use in an HIV clinic in Uganda	A pre–post study design using time-motion techniques and surveys. The authors found that providers who used the summary sheet were both able to spend more time directly with their patients and the average length of visit was reduced by 11.5 min.	Yes (each patient visit)	A largely successful process outcome. Explores the utility of summaries in a low-resource setting.
TimeLine/ AdaptEHR[40,41]	Abstraction from text and extraction of clinical variables, informative	Real structured, un-structured and image data on brain tumor patients	An interactive data display that summarizes and integrates various pieces of the EHR including images and free text.	A pilot study on Timeline found that although the initial learning curve was high, with time, the clinicians were able to perform image review quicker and were more confident in their clinical conclusions than when they used the EHR display.	No	Timeline had manually coded rules while AdaptEHR aims to automatically infer rules and relationships from ontologies and graphical models, the publication states that the conditional probability tables are not yet defined. Has four dimensions of representing data: time, space (where physical location of tumor), existence (certainty), and causality (treatment response treatment)
HARVEST[42]	Extraction of concepts from text and clinical variables, indicative	Real structured and un-structured EHR data	A problem-based, interactive, temporal visualization of a longitudinal patient record.	A task-based, timed evaluation found no difference in ability to extract, compare, synthesize and recall clinical information when using HARVEST in addition to the EHR, when carried out with subjects who had no prior experience with the summarization tool.	Yes (Real time)	Aggregates information from multiple care settings Operates on top of a commercial EHR system using HL7 messages Distributed computing infrastructure to enable real-time summarization.

The inputs, outputs, methods, and evaluation strategies are listed along with notable additional information for each summarizer.

REVIEW

identifying redundant sections along with language modeling techniques for assigning probabilistic similarity scores for phrase pairs.[45]

Concept-level Similarity

Concept-level similarity represents a more abstract level of similarity than similarity between words and strings. For instance, the concepts "epilepsy" and "seizure" – despite being two different UMLS concepts – share much semantic similarity when conveyed in a patient record.

In certain well-defined domains, clinical summarization approaches have relied on aggregating concepts, helping further the goal of synthesis[36,50] primarily through well-defined ontologies. For broader domains, how to identify that two semantic concepts are similar enough to be aggregated remains an open question. Furthermore, in text processing, mapping from words to concepts remains difficult because of the strong ambiguity of language.[43]

Detection of semantic redundancy has been investigated through two approaches: knowledge-free and knowledge-based. Knowledge-free similarity metrics have been developed for textual input. They rely on Harris' 1968 hypothesis which stipulates that concepts that appear in similar contexts are similar.[51] In practice, concepts are compared in a vector space, where each concept is a vector representing the context in which the concept typically occurs. This method has been implemented multiple times in the clinical domain to identify similar UMLS concepts.[52–54] Knowledge-free approaches are attractive when there is little ontological knowledge available. Alternatively, knowledge-based methods leverage existing resources to determine the similarity of two concepts. For instance, if the two concepts are present in an ontology, similarity can be assessed through the structure of the ontology. Other knowledge-based methods include examining similarity of the two concepts' definitions. We refer the reader to detailed reviews of concept-based similarity.[52,55] Despite the active research on this topic, these concept-level similarity methods have not been yet translated to most clinical summarization systems.

Statement-Level Similarity

A pervasive aspect of a patient record is the high level of statement redundancy across notes. For instance, two pathology reports for a given patient share many similar statements. Beyond the formulaic nature of documentation, statement-level redundancy also occurs because of copying and pasting from previous notes with some minimal editing of the copied statements.

In clinical summarization, there has been little work on this important aspect of similarity identification. Recently, a topic modeling approach was proposed to identify and control for such redundancy across patient notes.[56] In the general NLP community, identifying statement level similarity has been studied through the tasks of paraphrasing identification and textual entailment.[57] Many of the methods in text summarization for identifying both unidirectional (textual entailment) and bidirectional (paraphrasing) similarity employ a hybrid of methods for word-level and concept-level redundancy such as string similarity, logic-based methods, and context-vector.[58]

Along with the need for higher-order language similarity work in the clinical domain, there is an ongoing push to personalize similarity detection. It is well established that semantic similarity is context-dependent[59] and a recent study suggests that redundancy be examined as a function of the patient's previous history.[1] While identification of similar contexts based on the patient's health is an ongoing direction of research,[54] there is further work to be done in identifying context-specific similarity on higher-order semantic levels. Identifying similar words, concepts, and removing redundancy by patient-tailored

information aggregation is an important direction for future EHR summarization methodology.

2) ORGANIZING AND REASONING OVER TEMPORAL EVENTS

Patients' health evolves on many different time scales. Some health events such as pneumonia present themselves sporadically while chronic conditions like diabetes develop and worsen over a period of years. The importance of presenting clinical data in a time-dependent fashion has been recognized for a long time[60–62] however accurate temporal representation remains an open problem.[63–65] Automatic creation of a clinical data timeline from textual and structured clinical records requires temporal event extraction, ordering, and reasoning.

Temporality is an active research area in the genre of news summarization given the quick news cycle and fast-paced evolution of news stories.[66] However, news summarization research cannot always be readily translated into the health domain, as the challenges in health data are unique.[67,68] For example, different note types and specialties have different temporal relationships: pathology reports are often about one moment in time without reference to historical ailments whereas discharge summaries describe an entire inpatient hospital stay and instructions for future care. Styler et al. identified four complexities with extracting temporal information in clinical data: (i) diversity of time expressions; (ii) complexity of determining temporal relations among events; (iii) the difficulty of handling the temporal granularity of an event; and (iv) general NLP issues.[69]

After the extraction of event time, there is a need for performing relative temporal ordering.[70] Event ordering is difficult in part due to inexact wording, but also because clinical knowledge is often needed to infer how long conditions may last (e.g., a diabetes diagnosis is often not discussed at every visit but a clinician is aware that diabetes is a chronic condition, not an intermittently reoccurring condition each time the "diabetes" term is mentioned or the diabetes ICD-9 code is recorded).[71] Some recent work in event ordering includes the representation of temporal disease progression separately for each problem by Sonnenberg et al., an approach they call "clinical threading"[72] and frame-like semantic representations with rule-based temporal extraction to arrange problems on a timeline.[73] Raghavan et al.[74] identify and temporally order cross-narrative medical events across documents in clinical text using weighted finite state transducers.

Reasoning and abstraction of extracted clinical events to highlight disease progressions and trends is critical for creating succinct clinical summaries. Abstractions of temporal data can include combining events within a certain time frame and performing interval-based abstractions such as combining multiple chemotherapy drug mentions into a chemotherapy regimen time span[75] or reasoning about the length of time that symptoms lasted and their relation to diagnosis.[76] The questions of which events should be combined and what an appropriate time frame is remain difficult and currently resolved by leveraging clinical knowledge and ontologies. Time-dependent clinical summarization is a continuously evolving research area and there is opportunity for automatically identifying, accurately ordering, and performing reasoning over temporal clinical events.

3) ACCOUNTING FOR AND INTEPRETING MISSING DATA

Clinical records are sparse: documentation only occurs when a patient is seen by a clinician, thus clinical records miss the overwhelmingly large amount of observations about a patient across their lifetime. When summarizing sparse data, a critical complication is how to

interpret and reason over the missing data. In some cases, missing data is not important and can safely be ignored by a summarization system (e.g., a patient has no change in health status in between visits). In other cases, the presence of missing data hints at a salient aspect about the patient that needs to be highlighted within the summary (e.g., patient is too sick to come to their visit). How to interpret and determine the salience of missing data is a challenge, and one not investigated thus far in clinical summarization.

In the field of general statistics, there are three types of missing data: Missing Completely at Random, Missing at Random, and Missing Not at Random.[77] Most techniques for dealing with missing data assume that data are Missing Completely at Random or Missing at Random distributed, and include (i) variations of complete-case analysis, where only data with no missing values are used, (ii) single imputation, where missing data are imputed based on the values observed (using the mean, median, linear interpolation, etc.), and (iii) likelihood-based methods which compute maximum likelihood estimates for missing data.[78]

In the clinical domain, there is mounting evidence that most of the data are Missing Not at Random.[79,80] For these data, the missingness is informative, meaning that there is an underlying reason that the data are missing but that this reason is simply unobserved. Some techniques that use informative missing data properties to infer properties about clinical data have been proposed. A common way of using missing data in the clinical domain has been to look at how long values should last based on recorded measurements or documentation frequency. For example, laboratory test measurements have been studied to gather appropriate imputation time[81] and to infer health status features.[82] Van Vleck studied duration and persistence of problems in notes[83] as a function of missing data, while Klann[84] and Perotte[85] both studied the duration of ICD-9 codes. Klann estimated the durations for which each ICD-9 code remains valid and Perotte automatically classified ICD-9 codes into chronic and acute conditions. The modeling work that most explicitly demonstrates informativeness in missing data examined the accuracy of prediction models when: (i) ignoring missing data, (ii) interpolating missing data or (iii) incorporating a missing data indicator, and reported that the missing data indicator method performed best.[79] To properly provide context and infer trend lines, as demonstrated by Poh and de Lusignan for kidney disease data,[86,87] or to make predictions in clinical summaries it is critical to incorporate missing data literature and techniques into summarizer applications. The utility of modeling missing data explicitly is clear, however this conclusion is not being translated into clinical summarization research yet.

4) REDUCING INFORMATION TO ONLY THE MOST SALIENT

Salience identification has been heavily researched in the general domain text summarization literature. Early methods for identifying important topics relied on counts: frequency[88] and term frequency-inverse document frequency, which corrects for word specificity.[89] Other methods have focused on structure, such as document structure[90] or syntax structure[91] to identify important phrases. Syntactic information gleaned from the input document can identify which parts of a sentence are salient and which may be safely removed from a summary (e.g., a relative clause). It is unclear, however, how these approaches translate to the clinical domain, where syntactic structure is unconventional. Using prior knowledge of the input document structure (e.g., biomedical papers have an introduction, followed by a methods section) to weigh the salience of information pieces based on where they are conveyed in the document is, however, promising in

the clinical domain (yet not investigated thus far). Clinical notes follow a pre-specified structure; a diagnosis mention might be more relevant when conveyed in the past medical history than in the family history for instance. A different method for salience identification, still within the general domain summarization field, leverages discourse by considering sentences in input documents through a network, where lexical similarity between sentences is represented by the network edges. In this representation, salient sentences are the ones with the highest centralities.[92,93]

An alternative method for identifying relevant information relies on probabilistic modeling techniques such as Hidden Markov Models for identifying topics and topic changes in a set of documents[94] or hierarchical Latent Dirichlet Allocation-type models for identifying novel information with respect to older documents.[95] These Bayesian learning techniques for constructing effective automated summaries have also yet to be explicitly translated into the clinical arena.

The one type of salience detection that has been explicitly studied in the clinical domain is based on cue phrases. Cue phrases are pieces of text that signify that what follows is likely to be important. For example, "In conclusion" often precedes an important summarizing statement.[90] In clinical documentation, de Estrada et al.[96] developed a system called Puya that found cue phrases indicating normality or abnormality in the physical exam sections of notes. Another way of detecting salience relies on n-gram language modeling to identify the most recent information in the record, under the assumption that the newest information is the most salient for the provider to see.[97,98] A visualization prototype used this *n*-gram model to automatically highlight text that was found to be novel, drawing the provider's attention to the new findings.[99]

Defining salience in an operative fashion for automated summarization is an open question. In the general domain, there is evidence that humans sometimes disagree about what pieces of information are indeed salient, and that salience is often task-specific.[100] Similarly, in the clinical domain, determining what is important for a clinician is also probably quite task-specific. Nevertheless, it is safe to say that salience of elements in the patient record is related to capturing the health status of the patient and how it changes through time.[1,101] How to do so automatically, that is how to link textual and individual raw low-granularity observations to high-level clinical abstractions is one of the paramount challenge of informatics research. For instance, there has been little formal investigation of clinically specific markers of importance such as absolute change of a laboratory test value, the rate of change, the rate of mention of a particular concept, and other importance cues.

5) USING EXISTING CLINICAL KNOWLEDGE

The informatics community has invested enormous effort into codifying clinical knowledge in a variety of terminologies and ontologies. This knowledge representation effort has been successful in helping efforts like phenotyping combine terminological knowledge, expert reasoning, and machine learning to create actionable disease definitions.[102] Similarly in summarization work, it is important to make use of these available clinical knowledge representations and use them to generate rules and heuristics.

Several holistic summarization efforts leveraged terminologies to identify concepts that are semantically related (e.g., medications that treat particular conditions)[25] or rules to determine salience (e.g., identify and highlight the salient results that are abnormal).[30] However, summarization engines built for particular diseases benefit most often from manually crafted rules and disease-specific knowledge bases as

they enable tailored, task-dependent systems. The KNAVE-II application,[36] created for synthesis of bone marrow transplant patients, relies on an expert-maintained knowledge base for creating a semantic navigation system and concept abstraction. The Timeline system[40] is also built on a manually coded set of rules which identify salient concepts for different diseases, and perform temporal event reasoning. In addition, summaries that are setting and user specific often use expert-driven rules to ascertain which pieces of data should be shown at which time and to whom. Although the incorporation of clinical expertise into summarization is often a laborious process and sometimes only covers specific domains of expertise, it provides critical help in addressing some of the similarity, temporality and salience challenges. Of relevance to this review, we note that while existing summarizers rely on established knowledge resources, there is an active field of research to create these resources either by translating clinical expertise or acquiring the resources from data.[103–105]

6) DEPLOYING SUMMARIZATION TOOLS INTO THE CLINIC

The ultimate goal of any clinical summarization tool is implementation and usage by clinicians at the point of care. To date, however, there has been no widespread adoption of automated summarizers, especially for the large holistic temporal summarizers.[62] Pervasive deployment is often hindered by the commercial EHRs systems that have been adopted across the country. Building real-time computational tools to work atop commercially built EHR systems is still a daunting task as these vendor EHR systems are often not built to support interaction with outside applications. In addition, as the systems are closed off, dissemination of summaries across different hospitals and EHRs is a challenge as well. However, there is promising work with the i2b2-SMART platform that enables easier translation across institutions; researchers have developed a system to automatically link different data types across the EHR (mainly diseases and medications) and display a newly organized view of the patient record.[25]

To create meaningful and practical summaries that assist clinicians during their point of care needs, summarizers need to provide real-time information with patient record updates immediately available in the summary. This is an especially difficult task when the summary tool works with natural language, as the processing must be completed quickly and accurately. Current work with distributed infrastructures, like Apache Hadoop, provides promising results for immediate summarization.[42]

Another large barrier to translation of summarizer research into the clinical domain is rigorous evaluation. Hospitals often call for evidence of a useful summarizer before investing expensive resources into the implementation of the summarizer, but without adoption a summarizer is extremely difficult to evaluate. As is clear from Table 1, clinical summarization literature lacks standard evaluation metrics and there are very few extrinsic evaluations, a similar finding to a review of biomedical literature summarization by Mishra et al.[18] Given the restriction of limited adoption, it is not clear on which dimensions clinical summarizers should be evaluated. Initially, in order to avoid costly development and implementations with marginal benefit, it is imperative to study the need for a summarizer tool, context of usage, and clinician workflow. However, without eventual implementation into clinical care, showing any process- or health-level outcomes is not possible and therefore how to perform useful evaluations remains unclear: should, for instance, summarization systems focus on accurate information extraction, facilitating information exploration (e.g., which concepts are most relevant to the clinician), or user-friendly designs? Although the rigorous user-interface and cognitive process evaluations that are

necessary for creating new summarization systems often require deployment and study of actual use in practice, there exists guidance in the literature on cognitive aspects of clinical reasoning that can inform summarization system creation. Prior work on general medical cognition,[106] clinical decision-making,[107,108] human-computer interaction for interface design,[109–111] handoff communication,[112,113] clinical workflow analysis,[114,115] and some recent qualitative work specifically on clinical document synthesis which has identified common cognitive pathways for EHR document synthesis[1] and patterns of EHR data access[116] can guide the development of summarization systems. However, we emphasize that without actually studying the clinical context and manner in which clinicians use summarizers (either in the laboratory with prototype systems or in the clinic with deployed systems), it will be challenging to develop better evaluation strategies and better summarizers.

CONCLUSION

Within the past decade, the number of health practices that have some electronic capability to store patient data has grown to almost 80%. Health information exchanges promise patient record integration across multiple care settings and the amount of available patient data continues to explode.[117] The informatics community is posed to develop methods to mine the available information and ask questions such as: how can we further clinical knowledge, how can we assist clinicians in performing searches within and across patient records, how can we predict patient hospital course, and how can we automatically condense records to provide succinct summaries of a patient's medical history? With this eruption of rich, complex, and essential health data for millions of patients, the informatics community has new opportunity to tackle challenges of interpreting a mounting wealth of health information.

COMPETING INTERESTS

None.

CONTRIBUTORS

R.P. completed the literature review. R.P. and N.E. both identified existing gaps in the literature. R.P. and N.E. wrote the paper.

REFERENCES

1. Farri O, Pieckiewicz DS, Rahman AS, *et al*. A qualitative analysis of EHR clinical document synthesis by clinicians. *AMIA Annu Symp Proc*. 2012; 2012:1211–1220.
2. McDonald CJ. Protocol-based computer reminders, the quality of care and the non-perfectability of man. *N Engl J Med*. 1976;295:1351.
3. McDonald CJ, Callaghan FM, Weissman A, *et al*. Use of internist's free time by ambulatory care electronic medical record systems. *JAMA Intern Med*. 2014. Published online September 8, 2014, doi:10.1001/jamainternmed.2014.4506.
4. Holden RJ. Cognitive performance-altering effects of electronic medical records: An application of the human factors paradigm for patient safety. *Cogn Technol Work Online*. 2011;13:11–29.

REVIEW

5. Stead WW, Lin HS, eds. *Computational Technology for Effective Health Care: Immediate Steps and Strategic Directions.* Washington DC: National Academies Press; 2009.

6. Christensen T, Grimsmo A. Instant availability of patient records, but diminished availability of patient information: a multi-method study of GP's use of electronic patient records. *BMC Med Inform Decis Mak.* 2008;8:12.

7. Schiff GD, Bates DW. Can electronic clinical documentation help prevent diagnostic errors? *N Engl J Med.* 2010;362:1066–1069.

8. Laxmisan A, McCoy AB, Wright A, *et al.* Clinical summarization capabilities of commercially-available and internally-developed electronic health records. *Appl Clin Inform.* 2012;3:80–93.

9. Van Vleck TT, Wilcox A, Stetson PD, *et al.* Content and structure of clinical problem lists: a corpus analysis. *AMIA Annu Symp Proc.* 2008;2008: 753–757.

10. Rosenbloom ST, Shultz AW. Managing the flood of codes: maintaining patient problem lists in the era of meaningful use and ICD10. *AMIA Annu Symp Proc.* 2012;2012:8–10.

11. Powsner SM, Tufte ER. Graphical summary of patient status. *The Lancet.* 1994;344:386–389.

12. Payne TH. Computer decision support systems. *Chest.* 2000;118:47S–52S.

13. Feblowitz JC, Wright A, Singh H, *et al.* Summarization of clinical information: a conceptual model. *J Biomed Inform.* 2011;44:688–699.

14. Alterman R. Understanding and summarization. *Artif Intell Rev.* 1991;5: 239–254.

15. Radev DR, Hovy E, McKeown K. Introduction to the special issue on summarization. *Comput Linguist.* 2002;28:399–408.

16. Nenkova A, McKeown K. A survey of text summarization techniques. *Chapter in Mining Text Data.* 2012;43–76.

17. Afantenos S, Karkaletsis V, Stamatopoulos P. Summarization from medical documents: a survey. *Artif Intell Med.* 2005;33:157–177.

18. Mishra R, Bian J, Fiszman M, *et al.* Text summarization in the biomedical domain: a systematic review of recent research. *J Biomed Inform.* 2014. Published online July 10, 2014, doi:10.1016/j.jbi.2014.06.009.

19. Roque F, Slaughter L, Tkatsenko A. A comparison of several key information visualization systems for secondary use of electronic health record content. *In: Proceedings of NAACL HLT Workshop on Text and Data Mining of Health Documents*; 2010: 1–8.

20. Rind A, Wang TD, Aigner W, *et al.* Interactive information visualization to explore and query electronic health records: a systematic review. *Foundations Trends Hum-Comput Interact.* 2013;5:207–298.

21. West VL, Borland D, Hammond WE. Innovative information visualization of electronic health record data: a systematic review. *J Am Med Inform Assoc 2014.* Published online October 21, 2014, doi:10.1136/amiajnl-2014-002955.

22. Rogers JL, Haring OM. The impact of a computerized medical record summary system on incidence and length of hospitalization. *Med Care.* 1979; 17:618–630.

23. Liu H, Friedman C. CliniViewer: a tool for viewing electronic medical records based on natural language processing and XML. *Stud Health Technol Inform.* 2004;107:639–643.

24. Cao H, Markatou M, Melton GB, *et al.* Mining a clinical data warehouse to discover disease-finding associations using co-occurrence statistics. *AMIA Annu Symp Proc.* 2005;2005:106–110.

25. Klann JG, McCoy AB, Wright A, *et al.* Health care transformation through collaboration on open-source informatics projects: integrating a medical applications platform, research data repository, and patient summarization. *Interact J Med Res.* 2013;2:e11.

26. Rogers JL, Haring OM, Watson RA. Automating the medical record: emerging issues. *Proc Annu Symp Comput Appl Med Care.* 1979;3: 255–263.

27. O'Keefe QW, Simborg DW. Summary Time Oriented Record (STOR). *Proc 4th Ann Symp on Comp Appl in Med Care.* 1980;2:1175.

28. Powsner SM, Tufte ER. Summarizing clinical psychiatric data. *Psychiatr Serv Wash DC.* 1997;48:1458–61.

29. Plaisant C, Milash B, Rose A, *et al.* LifeLines: visualizing personal histories. *In: SIGCHI Conference on Human Factors in Computing Systems Proceedings*; 1996: 221–227.

30. Plaisant C, Mushlin R, Snyder A, *et al.* LifeLines: using visualization to enhance navigation and analysis of patient records. *Proc AMIA Annu Symp.* 1998;1998:76–80.

31. Friedman C, Alderson PO, Austin JH, *et al.* A general natural-language text processor for clinical radiology. *J Am Med Inform Assoc.* 1994;1:161–174.

32. Wilcox AB, Jones SS, Dorr DA, *et al.* Use and impact of a computer-generated patient summary worksheet for primary care. *AMIA Annu Symp Proc.* 2005;2005:824–828.

33. Hallett C, Scott D. Structural variation in generated health reports. *In: proceedings of the 3rd international workshop on paraphrasing*; 2005: 1–8.

34. Rogers J, Puleston C, Rector A. The CLEF chronicle: patient histories derived from electronic health records. *In: proceedings of the 22nd international conference on data engineering workshops*; 2006: 109.

35. Hallett C. Multi-modal presentation of medical histories. *In: Proceedings of the 13th international conference on intelligent user interfaces*; 2008: 80–89.

36. Shahar Y, Goren-Bar D, Boaz D, *et al.* Distributed, intelligent, interactive visualization and exploration of time-oriented clinical data and their abstractions. *Artif Intell Med.* 2006;38:115–135.

37. Hunter J, Freer Y, Gatt A, *et al.* Summarising complex ICU data in natural language. *AMIA Annu Symp Proc.* 2008;323–327.

38. Van der Meulen M, Logie RH, Freer Y, *et al.* When a graph is poorer than 100 words: A comparison of computerised natural language generation, human generated descriptions and graphical displays in neonatal intensive care. *Appl Cogn Psychol.* 2010;24:77–89.

39. Were MC, Shen C, Bwana M, *et al.* Creation and evaluation of EMR-based paper clinical summaries to support HIV-care in Uganda, Africa. *Int J Med Inf.* 2010;79:90–96.

40. Bui AAT, Aberle DR, Kangarloo H. TimeLine: visualizing Integrated Patient Records. *IEEE Trans Inf Technol Biomed.* 2007;11:462–473.

41. Bashyam V, Hsu W, Watt E, *et al.* Informatics in radiology: problem-centric organization and visualization of patient imaging and clinical data. *Radiographics.* 2009;29:331–343.

42. Hirsch J, Tanenbaum J, Lipsky Gorman S, *et al.* HARVEST, a longitudinal patient record summarizer. *J Am Med Inform Assoc.* 2014;22:263–274.

43. Friedman C, Elhadad N. Natural language processing in health care and biomedicine. In: Biomedical Informatics. Computer Applications in Healthcare. *Springer Science & Business Media,* New York, NY; 2014: 255–284.

44. Lindberg DA, Humphreys BL, McCray AT. The unified medical language system. *Methods Inf Med.* 1993;32:281–291.

45. Zhang R, Pakhomov S, McInnes BT, *et al.* Evaluating measures of redundancy in clinical texts. *AMIA Annu Symp Proc.* 2011;2011: 1612–1620.

46. Hirschtick RE. Copy-and-paste. *JAMA.* 2006;295:2335–2336.

47. Thornton JD, Schold JD, Venkateshaiah L, *et al.* Prevalence of copied information by attendings and residents in critical care progress notes. *Crit Care Med.* 2013;41:382–388.

48. Wrenn JO, Stein DM, Bakken S, *et al.* Quantifying clinical narrative redundancy in an electronic health record. *JAMIA.* 2010;17:49–53.

49. Cohen R, Elhadad M, Elhadad N. Redundancy in electronic health record corpora: analysis, impact on text mining performance and mitigation strategies. *BMC Bioinformatics.* 2013;14:10.

50. Hsu W, Taira RK, El-Saden S, *et al.* Context-based electronic health record: toward patient specific healthcare. *IEEE Trans Inf Technol Biomed.* 2012; 16:228–234.

51. Harris ZS. Mathematical Structures of Language. *Krieger Pub Co,* Melbourne, Florida, USA; 1968.

52. Pedersen T, Pakhomov S, Patwardhan S, *et al.* Measures of semantic similarity and relatedness in the biomedical domain. *J Biomed Inform.* 2007; 40:288–299.

53. Patwardhan S, Pedersen T. Using WordNet-based context vectors to estimate the semantic relatedness of concepts. *In: Proceedings of the EACL 2006 workshop making sense of sense*; 2006:1.

54. Pivovarov R, Elhadad N. A hybrid knowledge-based and data-driven approach to identifying semantically similar concepts. *J Biomed Inform.* 2012;45:471–481.

REVIEW

945

55. Pesquita C, Faria D, Falcão AO, *et al.* Semantic similarity in biomedical ontologies. *PLoS Comput Biol.* 2009;5:e1000443.

56. Cohen R, Aviram I, Elhadad M, *et al.* Redundancy-aware topic modeling for patient record notes. *PLoS One.* 2014;9:e87555.

57. Androutsopoulos I, Malakasiotis P. A survey of paraphrasing and textual entailment methods. *J Artif Intell Res.* 2010;38:135–187.

58. Dagan I, Dolan B, Magnini B, *et al.* Recognizing textual entailment: rational, evaluation and approaches–erratum. *Nat Lang Eng.* 2010;16:105.

59. Janowicz K. Kinds of contexts and their impact on semantic similarity measurement. *Sixth IEEE Int Conf on Perv Comp and Comm.* 2008;2008: 441–446.

60. Fries JF. Alternatives in medical record formats. *Med Care.* 1974;12: 871–881.

61. Cousins SB, Kahn MG. The visual display of temporal information. *Artif Intell Med.* 1991;3:341–57.

62. Samal L, Wright A, Wong BT, *et al.* Leveraging electronic health records to support chronic disease management: the need for temporal data views. *Inform Prim Care.* 2011;19:65–74.

63. Zhou L, Hripcsak G. Temporal reasoning with medical data–a review with emphasis on medical natural language processing. *J Biomed Inform.* 2007;40:183–202.

64. Sun W, Rumshisky A, Uzuner Ö. Temporal reasoning over clinical text: the state of the art. *J Am Med Inform Assoc.* 2013;20:814–819.

65. Wu ST, Juhn YJ, Sohn S, *et al.* Patient-level temporal aggregation for text-based asthma status ascertainment. *J Am Med Inform Assoc.* 2014;21: 876–884.

66. Allan J, Gupta R, Khandelwal V. Temporal summaries of new topics. *SIGIR.* 2001;2001:10–18.

67. Combi C, Shahar Y. Temporal reasoning and temporal data maintenance in medicine: issues and challenges. *Comput Biol Med.* 1997;27: 353–368.

68. Cios KJ, Moore GW. Uniqueness of medical data mining. *Artif Intell Med.* 2002;26:1–24.

69. Styler W, Bethard S, Finan S, *et al.* Temporal Annotation in the Clinical Domain. *Trans Assoc Comput Linguist.* 2014;2:143–154.

70. Savova G, Bethard S, Styler W, *et al.* Towards temporal relation discovery from the clinical narrative. *AMIA Annu Symp Proc.* 2009;2009:568–572.

71. Hripcsak G, Elhadad N, Chen Y-H, *et al.* Using empiric semantic correlation to interpret temporal assertions in clinical texts. *J Am Med Inform Assoc.* 2009;16:220–227.

72. Sonnenberg FA, Liu B, Feinberg JE, *et al.* Clinical threading: problem-oriented visual summaries of clinical data. *AMIA Annu Symp Proc.* 2012;353: 2433–2441.

73. Jung H, Allen J, Blaylock N, *et al.* Building timelines from narrative clinical records: initial results based-on deep natural language understanding. *Proceedings of BioNLP.* 2011;2011:146–154.

74. Raghavan P, Fosler-Lussier E, Elhadad N, *et al.* Cross-narrative temporal ordering of medical events. *ACL.* 2014;2014:998–1008.

75. Klimov D, Shahar Y, Taieb-Maimon M. Intelligent visualization and exploration of time-oriented data of multiple patients. *Artif Intell Med.* 2010;49: 11–31.

76. Zhou L, Parsons S, Hripcsak G. The evaluation of a temporal reasoning system in processing clinical discharge summaries. *J Am Med Inform Assoc.* 2008;15:99.

77. Little RJA, Rubin DB. *Statistical Analysis with Missing Data,* 2nd edn. New York, NY: John Wiley; 2002.

78. Enders CK. A primer on the use of modern missing-data methods in psychosomatic medicine research. *Psychosom Med.* 2006;68:427–436.

79. Lin J-H, Haug PJ. Exploiting missing clinical data in Bayesian network modeling for predicting medical problems. *J Biomed Inform.* 2008;41: 1–14.

80. Pivovarov R, Albers DJ, Sepulveda JL, *et al.* Identifying and mitigating biases in EHR laboratory tests. *J Biomed Inform.* 2014;51:24–34.

81. Hug CW. Predicting the Risk and Trajectory of Intensive Care Patients Using Survival Models. *Massachusetts Institute of Techonology,* Boston MA, USA; 2006.

82. Weber GM, Kohane IS. Extracting physician group intelligence from electronic health records to support evidence based medicine. *PLoS ONE.* 2013;8:e64933.

83. Van Vleck TT, Elhadad N. Corpus-based problem selection for EHR note summarization. *AMIA Annu Symp Proc.* 2010;2010:817–821.

84. Klann JG, Schadow G. Modeling the information-value decay of medical problems for problem list maintenance. *ACM IHI.* 2010;2010: 371–375.

85. Perotte A, Hripcsak G. Temporal properties of diagnosis code time series in aggregate. *IEEE J Biomed Heal Inform.* 2013;17:477–483.

86. Poh N, de Lusignan S. Modeling Rate of Change in Renal Function for Individual Patients: A Longitudinal Model Based on Routinely Collected Data. (NIPS PM 2011), Sierra Nevada. http://videolectures.net/nipsworkshops2011_poh_patients/.

87. Poh N, de Lusignan S. Data-modelling and visualisation in chronic kidney disease (CKD): a step towards personalised medicine. *Inform Prim Care.* 2011;19:57–63.

88. Luhn HP. The automatic creation of literature abstracts. *IBM J Res Dev.* 1958;2:159–165.

89. Jones KS. A statistical interpretation of term specificity and its application in retrieval. *J Doc.* 1972;28:11–21.

90. Edmundson HP. New methods in automatic extracting. *JACM.* 1969;16: 264–285.

91. Marcu D. From discourse structures to text summaries. *ACL.* 1997;97: 82–88.

92. Radev DR, Jing H, Budzikowska M. Centroid-based summarization of multiple documents: sentence extraction, utility-based evaluation, and user studies. *ANLP/NAACL Workshop on Summarization.* 2000;21–30.

93. Erkan G, Radev DR. LexRank: Graph-based lexical centrality as salience in text summarization. *J Artif Intell Res.* 2004;22:457–479.

94. Barzilay R, Lee L. Catching the drift: probabilistic content models, with applications to generation and summarization. *Proc HLT-NAACL.* 2004; 113–120.

95. Delort J-Y, Alfonseca E. DualSum: a topic-model based approach for update summarization. *ACL.* 2012:214–223.

96. De Estrada WD, Murphy S, Barnett GO. Puya: a method of attracting attention to relevant physical findings. *AMIA Annu Symp Proc.* 1997;1997: 509–513.

97. Zhang R, Pakhomov S, Melton GB. Automated identification of relevant new information in clinical narrative. *2nd ACM IGHIT Symp Proc.* 2012; 2012:837–842.

98. Zhang R, Pakhomov S, Melton G. Longitudinal analysis of new information types in clinical notes. *AMIA CRI.* 2014;2014:1–6.

99. Farri O, Rahman A, Monsen KA, *et al.* Impact of a prototype visualization tool for new information in EHR clinical documents. *Appl Clin Inform.* 2012; 3:404–418.

100. Nenkova A, Passonneau RJ. Evaluating content selection in summarization: the pyramid method. *Proc of HLT-NAACL.* 2004;4:145–152.

101. Suermondt HJ, Tang PC, Strong PC, *et al.* Automated identification of relevant patient information in a physician's workstation. *Proc Annu Symp Comput Appl Sic Med Care Symp Comput Appl Med Care.* 1993;1993: 229–232.

102. Pathak J, Kho AN, Denny JC. Electronic health records-driven phenotyping: challenges, recent advances, and perspectives. *J Am Med Inform Assoc.* 2013;20:e206–e211.

103. Noy NF, Shah NH, Whetzel PL, *et al.* BioPortal: ontologies and integrated data resources at the click of a mouse. *Nucleic Acids Res.* 2009;37: W170–W173.

104. Mortensen JM, Horridge M, Musen MA, *et al.* Applications of ontology design patterns in biomedical ontologies. *AMIA Annu Symp Proc.* 2012;2012: 643–652.

105. Tao C, Song D, Sharma D, *et al.* Semantator: semantic annotator for converting biomedical text to linked data. *J Biomed Inform.* 2013;46: 882–893.

106. Patel VL, Arocha JF, Kaufman DR. A primer on aspects of cognition for medical informatics. *AMIA Annu Symp Proc.* 2001;8:324–343.

REVIEW

107. Arocha JF, Wang D, Patel VL. Identifying reasoning strategies in medical decision making: A methodological guide. *J Biomed Inform.* 2005;38:154–171.

108. Kushniruk AW. Analysis of complex decision-making processes in health care: cognitive approaches to health informatics. *J Biomed Inform.* 2001; 34:365–376.

109. Patel VLV, Kushniruk AWA. Interface design for health care environments: the role of cognitive science. *AMIA Annu Symp Proc.* 1998;1998:29–37.

110. Jaspers MWM, Steen T, van den Bos C, *et al.* The think aloud method: a guide to user interface design. *Int J Med Inf.* 2004;73:781–795.

111. Thyvalikakath TP, Dziabiak MP, Johnson R, *et al.* Advancing cognitive engineering methods to support user interface design for electronic health records. *Int J Med Inf.* 2014;83:292–302.

112. Abraham J, Nguyen V, Almoosa KF, *et al.* Falling through the cracks: information breakdowns in critical care handoff communication. *AMIA Annu Symp Proc.* 2011;2011:28–37.

113. Abraham J, Kannampallil TG, Almoosa KF, *et al.* Comparative evaluation of the content and structure of communication using two handoff tools: implications for patient safety. *J Crit Care.* 2014;29:311.e1–7.

114. Unertl KM, Weinger MB, Johnson KB, *et al.* Describing and modeling workflow and information flow in chronic disease care. *J Am Med Inform Assoc.* 2009;16:826–836.

115. Militello LG, Arbuckle NB, Saleem JJ, *et al.* Sources of variation in primary care clinical workflow: implications for the design of cognitive support. *Health Informatics J.* 2014;20:35–49.

116. Reichert D, Kaufman D, Bloxham B, *et al.* Cognitive analysis of the summarization of longitudinal patient records. *AMIA Annu Symp Proc*; 2010;2010: 667–671.

117. Adler-Milstein J, Bates DW, Jha AK. A survey of health information exchange organizations in the United States: implications for meaningful use. *Ann Intern Med.* 2011;10:666–671.

AUTHOR AFFILIATIONS

Department of Biomedical Informatics, Columbia University, New York, USA

REVIEW