Published in final edited form as:
J Biomed Inform. 2012 December ; 45(6): 1151–1163. doi:10.1016/j.jbi.2012.07.012.

A classification of errors in lay comprehension of medical documents

Alla Keselman, MA, PhD[1] and **Catherine Arnott Smith, AMLS, MS, MA, PhD**[2]

[1]Division of Specialized Information Services, National Library of Medicine, Bethesda, MD, United States

[2]School of Library and Information Studies, University of Wisconsin-Madison, Madison, WI, United States

Abstract

Emphasis on participatory medicine requires that patients and consumers participate in tasks traditionally reserved for healthcare providers. This includes reading and comprehending medical documents, often but not necessarily in the context of interacting with Personal Health Records (PHRs). Research suggests that while giving patients access to medical documents has many benefits (e.g., improved patient-provider communication), lay people often have difficulty understanding medical information. Informatics can address the problem by developing tools that support comprehension; this requires in-depth understanding of the nature and causes of errors that lay people make when comprehending clinical documents. The objective of this study was to develop a classification scheme of comprehension errors, based on lay individuals' retellings of two documents containing clinical text: a description of a clinical trial and a typical office visit note. While not comprehensive, the scheme can serve as a foundation of further development of a taxonomy of patients' comprehension errors. Eighty participants, all healthy volunteers, read and retold two medical documents. A data-driven content analysis procedure was used to extract and classify retelling errors. The resulting hierarchical classification scheme contains nine categories and twenty-three subcategories. The most common error made by the participants involved incorrectly recalling brand names of medications. Other common errors included misunderstanding clinical concepts, misreporting the objective of a clinical research study and physician's findings during a patient's visit, and confusing and misspelling clinical terms. A combination of informatics support and health education is likely to improve the accuracy of lay comprehension of medical documents.

Keywords

Information Literacy; Clinical Documentation; Consumer Health; Patients; Classification; Content Analysis

Corresponding author: Alla Keselman, PhD, MA, Division of Specialized Information Services, National Library of Medicine, 6707 Democracy Blvd., Suite 510, Bethesda, MD 20892-5467, Phone: 301-496-3420, Fax: 1-301-480-3537, keselmana@mail.nih.gov.

1. BACKGROUND

1.1 Introduction

Today's emphasis on participatory medicine calls for patients to take an active role in their healthcare. This requires that lay individuals participate in tasks traditionally reserved for health care providers. In particular, patients and families are expected to interact with a large number of health and medical documents. While some of these documents, such as those found on patient education web sites and informed consent forms, are written specifically for lay health consumers, others, such as medical records, are not. There is a growing emphasis on Personal Health Records, which can contain electronic documents helping patients both to keep abreast of and to contribute to the information flow of the healthcare process. At the present time, there is no consensus as to what a PHR should contain, but it is expected that in the future most PHRs will provide access to fragments or whole documents authored by clinicians [1]. "Current conceptualizations of the personal health record," caution Brennan et al. "carry an implicit expectation that a person (clinician, patient, parent) must literally read, then process the specified content of the record."[2] While the usefulness of enabling patient access to medical documents depends on patients' ability to understand clinical content, literature offers little discussion of lay comprehension accuracy and errors in the context of such documents. The only exception is studies of patients' comprehension of informed consent, reviewed below.

This paper analyzes comprehension errors that lay people make when reading examples of two document types, a description of a clinical trial and a physician's visit note, and proposes a classification scheme for these errors. Understanding categories and causes of lay comprehension errors is essential for development of informatics support for the task and designing useful, usable PHRs.

1.2 Patients' experience with PHRs and EHRs

An ultimate tool in patient empowerment is a Personal Health Record (PHR), a tailored variant and document subset of the Electronic Health Record, with the patient as an intended co-creator and user of the content. Numerous commercial companies and research groups are developing PHR models and applications (e.g., Brennan [3]; Hampton [4]). The assumption is that participating in creating, managing and using their health information increases individuals' health knowledge and leads to "greater responsibility for their own health and well-being" [5]. Although this study does not specifically focus on PHR comprehension errors, the two documents it employs are representative of the content and level of complexity of PHR information. This suggests that 1) PHR research can help formulate the framework for studying lay errors in understanding medical documentation and 2) a study of lay comprehension of medical documents can make a contribution to the PHR discourse. The current PHR research agenda focuses on issues of architecture, attitudes and adoption among providers and patients, and related privacy and security concerns [6]. However, the problem of supporting patients in the potentially challenging task of co-authoring and using a professional medical document receives limited attention. This is despite the fact that most PHR models are highly complex. The HIMSS Minimum Data Set for PHRs recommends including data from clinical summaries (i.e., active and historical prescriptions; current OTC meds; allergy information, diagnoses; problem limits; immunization status); results and reports; and histories (immunization, past medical, past surgical, family and social) [7]). Marshall, of WebMD, adds to this list specialized types of medical images, such as X-rays and mammograms, as well as EKG readings and even DNR directives [8].

Physicians have raised concerns about the accuracy of patient-contributed PHR data; even among those willing to use PHRs, 71% said in one recent study that they were "somewhat or very" concerned that the PHR might contain incorrect information [9]. However, few studies provide insight about this problem, or the degree to which patients comprehend the information they have read that was authored by health professionals. Wuerdeman and colleagues [10] studied the accuracy of information that patients contributed to their Electronic Health Records. Patient-reported data about medical tests and results were compared against the data entered by health care providers into the EHR. When patients were asked whether a specific test had been performed for them, the match between their response and the providers' entry ranged from 78% (proctoscopic exam) to 90% (stool test). The range for test values was slightly lower, with matches as low as 70% for LDL (within 10 points) to 88% for total cholesterol (within 20 points). Depending on the reader's perspective, these findings may suggest that patients' recall of information is mostly accurate, or often inaccurate. However, these authors also find that patients are often able to provide valuable information that is not in the EHR, concluding that patient entry constitutes a viable, if not always accurate, source of information.

Kim and Johnson also evaluated patient accuracy in a different kind of study. Their patient participants manually entered information into PHRs that they extracted from transcripts with clinical content, such as clinic notes, medications lists, and laboratory test results [11]. This resulted in spelling errors, a surprising finding because these subjects were expected to have familiarity with the medical terminology used in their diagnostic domain. Besides these errors, Kim and Johnson found inconsistencies in the type of content that patients considered important to enter in their PHRs; these inconsistencies relate to the deeper challenges of scientific literacy. Specifically, the subjects in this study were found to have entered extraneous information – particularly in the free-text sections of the PHR – but also to have omitted quantitative information. Each of these problems has implications for data quality and data processing of PHRs. In another study conducted in the same setting, Kim et al. [12] concluded that "Low health literacy … was also an important factor that limited PHIMS [Personal Health Information Management System] use. Some users … commented that they preferred to use it with a nursing student who could provide explanations for them to understand their health information." Tran et al. also identified "problematic jargon" during prototype testing of a PHR [13], and Lober et al. found that overall health literacy, manifest by questions about conditions, medications, terminology, and more, presented a barrier to almost of third of their subjects [14]. Britto et al. [15] testing a pediatric patient portal for parents, reported participants who didn't recognize abbreviations ("Fe" for iron) and didn't know what a pathology report was. Jargon translations, medical interpretations and explanations, particularly for labs, were "often requested" by subjects.

The level of accuracy of patients' self-report of their medical history found by Wuerderman et al. [10] is consistent with that found in non-EHR-related studies of patients' self-report [16, 17, 18]. For example, Khoja and colleagues [17] found that patients' self-reported history of colorectal cancer screening was in good agreement with their physicians' reports. Collectively, these studies suggest that patient-generated information constitutes potentially very valuable, if not error-proof, contribution to the PHRs.

With regard to patients' use of the records authored by health professionals, research suggests some positive effect on doctor-patient communication and patient satisfaction (Ross and Lin [19]; Businger et al. [20]). Less is known about the particulars of patients' use of the records, as well as specific difficulties that patients may experience. Keselman and colleagues [21] surveyed 104 patients about their experience reviewing their health records (paper or electronic), including the ease of comprehension. These patients reported that the sections they found the easiest to understand were immunization records, medication lists

and discharge summaries (respectively rated as "easy" by 80%, 71% and 63% of the respondents who answered that particular question). The sections considered most difficult to understand were physicians' notes, radiology reports and nurses' notes (rated as "easy" by 36%, 45% and 47% of responders). Qualitative analysis of narrative comments provided some insight into barriers to successful use of records. These included, in the order of frequency: problems with records access, lack of conceptual knowledge, problems with medical language (e.g., terminology), poor data quality in the records, and difficulty dealing with the records' structure, organization, and lack of standardization. Similarly to Wuerdeman et al [10] this study suggests that patients' participation in records maintenance is important, and that problems with comprehension exist and need addressing.

Several authors note that despite potential benefits, patients' engagement with EHRs and PHRs is relatively low (for review, see Archer et al [22]). One barrier to adoption, identified in the research literature on PHRs, is terminology. Clinical content is written in clinical terminology. Decades of research focusing on encounters between patients and provider-held records have demonstrated that, for some patients, medical terminology presents an obstacle towards effective understanding of clinical content. As Lee et al. express it: "If users cannot understand the content, such contents are useless" (S 313) [23]. The comprehensive review of PHRs' evidence base compiled by Marchionini et al. [24] identifies terminology as one of three necessary types of "experience" for PHR users. Maloney and Wright further comment that patients who do not understand their condition do not produce complete records of their condition, although the physician may incorrectly *perceive* such records as complete [25]. Although the review presented in this section suggests that terminology is not the only barrier to patients' user of PHRs, as other aspects of health literacy and scientific literacy are often part of the problem, informatics' response has been primarily in the area of vocabulary support. Baorto and Cimino [26] developed an "infobutton" application that produces explanations of medical terms in Pap smear reports and links patients to free relevant web resources. Zeng-Treitler and colleagues [27] developed a prototype EHR translator which identifies and replaces medical terms that are difficult for consumers. Extra-terminological approaches for supporting patients include bundling PHRs with health education materials and providing tools for decision making / risk appraisal[28, 29]). For example, Adnan, Warren, and Orr [30] developed a tool that identifies difficult terms in discharge summaries and hyperlinks them to consumer-friendly MedlinePlus pages. Further research and development of interventions would benefit from additional research on patients' comprehension of EHR and PHR content.

1.3 Patients' experience with informed consent

As the informed consent document is the only type of medical text of which patients' comprehension has been studied extensively, studies of informed consent comprehension merit our special attention. Many of these works focused on consent to clinical trials; others studied consent to standard care procedures. Like other medical texts, informed consent forms typically score above the 6-8th grade readability level recommended for patient education documents. For example, after reviewing informed consent templates on IRB websites of 114 U.S. medical schools, Paasche-Orlow and colleagues (2003) [31] found that the average readability level was 10.6 according to the Flesch-Kincaid formula. This average actual readability level exceeded the typical IRB-recommended score by 2.8 grade levels. These high reading level scores can in part be explained by the conceptual complexity of the medical content of the informed consent document: describing complex medical procedures at the 8th grade level is challenging. In addition, the informed consent document suffers from an "identity crisis". On the one hand, its goal is to provide patients with clear, complete, accurate information about the treatment or trial in question, as well as the alternative care options. On the other hand, the informed consent is also a legal

document, conceived as the evidence in a potential litigation process, which makes it long and anything but simple.

As could be expected given these constraints, informed consent documents are difficult for patients to understand and remember. Joffe et al. [32] administered an informed consent understanding questionnaire to 207 adult cancer patients enrolled in a clinical trial. They found that 63% of the responders did not fully understand the risks involved and 70% did not understand the uncertain nature of the treatment. In contrast, participants' subjective perception of their level of understanding was high – 90% considered themselves well-informed. Findings of low recall and comprehension coupled with high satisfaction with the consent process are echoed in other studies. In a study comparing two informed consent formats, Olver et al [33]found that 49-51% of cancer patients receiving chemotherapy for the first time could not correctly recall the number of drugs they received, and 45-55% did not correctly recall the treatment goal. Most patients were satisfied with their understanding of chemotherapy, although less than half read all the information provided. Unfortunately, we were not able to find published studies that attempted to classify the broad range and causes of specific errors in informed consent comprehension. For example, while Joffe et al. [32] asked patients to agree or disagree with (accurate) statements about the clinical trials in which they participated, they did not probe for the specific misconceptions underlying disagreement. Similarly, while Olver et al. [33] mention that patients in their study were asked to name the chemotherapy drugs that they received, these authors do not provide details about which names were most likely to escape recall. Both studies also limited their scope to a few easily quantifiable variables.

Given all the difficulties in understanding informed consent, it is not surprising that a large number of studies have attempted to identify ways to improve consent documents. Unfortunately, the results are mixed at best. In 2004, Flory and Emanuel [34] published a systematic review of 42 clinical trials of interventions designed to improve "research participants' understanding of informed consent in research." Interventions described in this review fell into the following five categories (numbers in parentheses indicate total number of studies in that category, as well as the number of studies reporting significant comprehension improvement):

1. Multimedia interventions [12 studies]
2. Enhanced content, length, writing style or format [15 studies]
3. Extended discussion [5 studies]
4. Test/feedback [5 studies]
5. Miscellaneous [5 studies]

Significant improvement in understanding was demonstrated in only 3 out of 12 multimedia trials (25%), 6 out of 15 enhanced consent forms trials (40%), and 3 out of 5 extended discussion trials (60%), with the remaining two extended discussion trials showing trends towards improvement. All test/feedback trials (100%) showed significant improvement in comprehension, but according to Flory and Emanuel [34], it is possible that outcome measures of these studies tapped route memorization rather than genuine comprehension. At face value, it appears that simply editing the document or converting the text into a multimedia format does not work. Interventions with the greatest promise for improvement in patient comprehension are the ones most difficult to implement, because they require one-on-one discussion of the protocol between a potential participant and a trial team member. Moreover, even this level of individualized support may not produce genuine comprehension.

Attempts to improve patients' informed consent to standard, rather than experimental, treatments (e.g., non-experimental oral surgery, anesthesia, or chemotherapy) have similarly limited impact on recall and comprehension. For example, in a study mentioned earlier, Olver et al. [33] found that presenting chemotherapy information on a CD-ROM rather than on a written consent form "did not improve cancer patients' recall of treatment information enough to warrant changes in consent procedures." In a similar study in the orthodontics domain, Kang et al [35] found that improving readability of informed consent had little effect on its recall and comprehension, although the addition of a narrated Power Point slide show did produce improvement. However, just like the chemotherapy patients researched by Olver et al. [33], participants in the narrated slide show condition overestimated their understanding.

In summary, informed consent forms are often written above the 6-8th grade readability level recommended for consumer health materials. Studies of patients interacting with these informed consent forms suggest that many participants do not fully understand important aspects of clinical trials and standard care procedures, and overestimate their level of understanding. Attempts to increase understanding have mixed results. Improving the readability and the format of the text itself appears to be of limited effectiveness. Findings of the effectiveness of one-on-one discussions and multimedia interventions are somewhat ambiguous, due to the methodological flaws of the studies. While informed consent studies identify a variety of aspects of clinical trials that patients have difficulty comprehending, they do not provide systematic information about the causes and the specific nature of patients' comprehension errors.

1.4 Taxonomies of errors in healthcare

Literature reviewed in the previous sections suggests that errors in lay comprehension of medical documents are common. Classification of these errors would help health educators to understand their causes and ultimately, develop comprehension support tools. We were not able to find any published taxonomies or classification schemes or errors that patients / consumers make when interacting with health information. Medical errors literature provides numerous examples of classifying mistakes of health professionals. These classifications are usually not grounded in any theoretical framework. Instead, they serve as a basis for error reporting systems, and thus reflect reporting needs of specific areas of specialization and institutions. Some specializations, such as family practice, nursing, or primary care, are broad (Dovey [36]; Benner [37]; Ely [38]) and characterized by a reasonably extensive amount of classification work; others, such as pediatric prescribing, are narrow and not well developed (Davis [39]). In addition to varying according to the area of specialization, the classification schemes differ in the dimensions along which the classification is conducted.

In the broadest sense, there are two aspects of errors that are of interest: 1) things that go wrong, and 2) reasons why this happens [40]. Within each specialization, things that go wrong can be described in terms of the clinical procedure or stage of care (e.g., surgery, post-operative care), severity, physical location, and many other attributes. In recent years, there have been several efforts to standardize classification of medical errors, while making them comprehensive and theory-based. JCAHO proposed a taxonomy that classifies errors along five dimensions:

- Impact (e.g., psychological, physical, economic)
- Type (communication, patient management, clinical performance)
- Domain (physical setting, staff involved, patient characteristics)
- Cause (organizational, technical, human)

- prevention / mitigation [40]

The JCAHO classification is an attempt to develop a standardized format that can serve as a backbone onto which terminologies used by different electronic systems can be mapped. A classification of this type also provides a solid basis for studying and preventing medical errors. In a different attempt to provide theoretical basis for a study and prevention of medical errors, Zhang and colleagues [41] proposed a cognitive taxonomy, classifying errors along two dimensions. The first can be described as cause / intentionality: “slips”, or incorrect executions of correct steps, are contrasted with “mistakes”, or correct executions of incorrect steps. The second is the phase of the action process: execution slips occur during goal formulation and action specification; evaluation slips occur as results of actions are perceived and interpreted.

Review of taxonomies of medical errors reveals that classification systems depend on their purpose / function, as there are many possible and useful ways to partition the data. Beyond that, however, because these taxonomies derive from analysis of professional activity, they are not useful for classifying lay errors in processing medical documents. Some insight about the nature and types of difficulties that lay readers are likely to experience can be obtained from studies of health literacy. For example, Chan, Matthews, and Kaufman [42] propose a two-dimensional taxonomy of eHealth literacy, which classifies eHealth tasks according to 1) complexity of cognitive processes and 2) dimensions of literacy that are involved. Complexity is based on the level of cognitive effort; for example, applying and analyzing information is considered more complex than understanding it. Literacy dimensions involve computer, information, media, traditional literacy and numeracy, scientific, and health literacies. Of these, traditional literacy and numeracy, scientific and health literacy (which includes knowledge of medical terminology) are likely to be the dimensions most relevant to understanding lay errors in comprehending medical documents. Unfortunately, of these, scientific and health literacy themselves are concepts without agreed-upon, solid conceptual definitions. More work exists in the domain of health numeracy, with studies focusing on lay adults’ abilities to perform basic calculations, understand risk and statistics, comprehend different number formats (e.g., simple fractions vs. frequencies), and interpret different graphical representations of numerical data [43]).

1.6 Specific research objectives

Our intent was to develop a data-driven scheme, based on participants’ comprehension of two documents. These documents are representative of the types of medical texts that are not only notoriously challenging, but important for patients’ participation in their care. Just as the medical error taxonomies reported above rely on clinical activity and experience, our scheme is driven by the experience and activity of laypeople/consumers, interacting with real clinical documents. Research focused on supporting patients’ comprehension of medical documents, such as patient records or description of clinical trials, would benefit from a study of the specific nature of comprehension errors. The objective of this exploratory study was to develop a classification scheme of comprehension errors, based on lay individuals’ retellings of two medical documents: a description of a clinical trial and a cardiology office visit note. Developing a comprehensive taxonomy of patients’ and consumers’ errors in comprehending medical documents was beyond the scope of this work. This project was intended to 1) demonstrate that such development is feasible, 2) outline a set of methods suited to the task, and 3) develop a scheme that could provide a foundation for the taxonomy.

2 METHODS

2.1 Participants

The study involved 80 participants, all staff, faculty or undergraduate or graduate students at the University of Wisconsin-Madison. All participants were native English speakers. They were recruited via campus fliers and newspaper advertising (40 participants), or via an in-class announcement in a graduate course in library and information studies (40 participants). All participants received $25 bookstore gift cards for participation. The study was approved by the Social Sciences Institutional Review Board of the University of Wisconsin-Madison on February 23, 2007. All participants completed an anonymous questionnaire that included demographic questions (gender, age, racial/ethnic characteristics, educational level, and work experience) as well as self-assessment questions of biomedical understanding and knowledge about diabetes mellitus (Table 1). Subjects self-rated their biological knowledge on a scale on a scale from 1 (“I rarely read texts on biomedical topics”) through 4 (“I read and understand general medical articles”) and self-rated their knowledge about diabetes mellitus on a scale from 1 (“very little”) to 5 (“a good deal”). (see Table 1 legend).

2.2 Content Analysis

Content analysis has been defined as “A group of formal … techniques used to analyze texts” [44]. It is a method frequently used in studies where text is the data of interest, for example, patient-physician email and web-based messages in PHRs (for a review see [45]); Web-based and print-based consumer-authored content [46, 47]; the published literature of medicine [48]; and transcribed interviews [49].

In the research literature of medical records and clinical documentation, the focus of content analysis has most often been on material authored by health professionals and of which patients are the subjects, not the authors; for example, studies of contraceptive use and its presence in the medical records of teenagers requesting an abortion [50]; cultural background of patients in documentation [51]; and connections to end-of-life care documented in ICU forms [52].

Content analysis is also done in usability studies; for example, Haas et al.[53] researched physician users’ perceptions of the usability of an electronic clinical note. From the patient perspective, Arar et al.[54] interviewed veterans about their experiences using the Surgeon General’s online Family Health History tool. Ayana, Pound and Ebrahim used focus groups of therapists to gather opinions about the usefulness of a patient-held record for stroke patients [55].

Three studies were found in the medical literature that used content analysis as a tool to understand medical records from the patients’ or lay person’s viewpoint. In these studies, it was patients’ opinions which were of interest and not the records themselves. Wibe et al. [56] interviewed 17 patients to understand their feelings about requesting and reading a copy of their medical record. Content analysis was used to explore themes in their narrative. Bhavnani et al. used content analysis to analyze the responses to a questionnaire asking patients in the UK about their access to their electronic medical record [57]. Rassin et al. used content analysis to better understand 28 cardiac patients’ management of their medical documents [58].

There is precedent in the medical literature for using content analysis to inform development of taxonomies; for example, Saboor and Ammenwerth [59] and an error categorization for information and communication systems in hospitals, and Whitson et al. who worked with patients in rehabilitation clinics to create a taxonomy of comorbidities and their effects [60].

2.3 Procedure

Each participant in the study read two documents, a description of a clinical trial and a sample office visit note. The error analysis which is the focus of this paper (described below) was conducted in conjunction with an intervention study, reported elsewhere [61] (). The intervention had four conditions. Participants in Condition 1 read the documents in their original, unaltered form (see "Document Types" below); Participants in Conditions 2 and 3 read versions that included two different types of vocabulary support for difficult terms; Participants in Condition 4 read versions that were edited / rewritten to enhance the documents' coherence, or connectedness of ideas. The objective of the intervention was to achieve improvement in completeness (as measured by the number of sentence clauses recalled) and accuracy (as measured by the number of errors) of texts' recall in two of the four conditions. As the results of the intervention analysis demonstrated some expected improvement in recall completeness, but not in accuracy (the number of errors), the four conditions were combined for the current analysis. The rest of this section reports the procedure without making reference to participants' conditions; the findings are also reported in aggregate form across the conditions.

Participants worked on individual computers; a research assistant was present at all times, ensuring that work was done individually. The order of presentation of the Clinical Trial and Visit Note documents was randomized among participants. After completing the anonymous demographic questionnaire described above, participants read their first document on the computer screen. After a waiting period of 10 minutes, they wrote their recollection of the document's text using Microsoft Word. Participants were instructed to retell the document as if sharing its information with a person who had never seen it before. This procedure was then repeated for each subject's second document.

2.4 Documents[1]

The documents chosen for presentation to study participants were selected because they are representative of those that patients are likely to encounter when they participate in clinical trials and navigate "tethered" PHRs.

The first, the Clinical Trial (Textbox 1), was adapted from a record found at ClinicalTrials.gov, the largest existing database of clinical trials, maintained by the National Library of Medicine (database trial identification NCT00481598). One motivation behind creating ClinicalTrials.gov's database was the desire to make clinical trial information "available to individuals with serious or life-threatening diseases and conditions, *to other members of the public* [emphasis ours—Authors], to health care providers, and to researchers" and available "in a form that can be readily understood by members of the public." Clinical trial records often constitute patients' first encounter with the information about a medical study of potential interest, and are read without a health professional's assistance. This makes them appropriate for a study assessing consumer text comprehension. This particular trial record was selected because it involved a common diagnosis—diabetes--and included a description of the trial's purpose (Textbox 1).

The second document, a Visit Note (Textbox 2), was selected because it addresses several health concerns of general interest to the public. The scoping review of Archer et al. identifies "notes" as a data source recommended for inclusion in PHRs at the recommendation of the American Medical Informatics Association's College of Medical

[1]This section describes the original versions of each document. Modified documents presented to participants in Conditions 2-3 included the same content, but also vocabulary definitions that appeared in balloons over difficult words, as these words were moused over. Versions presented to participants in Condition 4 included some additional information, and in the case of the Visit Note, additional subsections within document sections.

Informatics [22]. The U.S. Military Health System piloted a PHR in partnership with two commercial products, Microsoft® HealthVault and Google® Health (now defunct); HealthVault allows for transfer of CCR and CCD standard documents from EHRs, and the Military Care PHR featured inpatient notes and outpatient encounter notes [62].

Like the Clinical Trial, it involves diabetes, but because the anonymous patient described has multiple problems, the Visit Note also incorporates cardiology. This Visit Note included five sections: a) History of present illness; b) Physical examination; c) Medications; d) Diagnoses, and e) Plan. Previous research into a large dataset of electronic medical record documents conducted by the second author validates that these five sections were found in 37-78% of notes in the University of Pittsburgh Medical Center's MARS system [63]. The Note was obtained from MedicalTranscriptionSamples.com, an online collection of transcripts for medical transcriptionist training. A nurse practitioner and a physician read the document and confirmed that it was representative of an office visit note.

3 CODING AND ANALYSIS

3.1 Coding procedure

Authors CAS and AK jointly read all retellings written by the participants, marking any errors that distorted the meaning of some statement in the original document, or that made a claim not supported by the original document. In this first round, the authors erred on the side of over-extraction, marking anything that could potentially constitute an error. They then conducted several joint rounds of review and discussion, turning to reference literature and consulting a clinician about contentious cases. Inferences, defined as statements that were not made in the original texts, but could be viewed as reasonable conclusions from the text, were not counted as errors. The final error list consisted of 157 Clinical Trial and 220 Visit Note errors. CAS and AK jointly developed the error classification scheme using a grounded theory approach [64]. The approach involved conducting several iterative reviews of the errors for each document type, while assigning data-driven descriptive labels and grouping and re-grouping related errors into categories. The procedure started with top-level codes (e.g., Findings) in the early iterations, with the development of sub-codes (e.g., Finding Inaccurately Reported, Finding Nonsensical) in later iterations. After the process was performed for each document type separately, a joint classification scheme was developed by merging the two (Table 2).

Next, AK and CAS jointly coded the retellings, resolving disagreements via discussion. The coding procedure allowed applying multiple codes to a single error statement (e.g., Finding, Incorrectly Reported could also be coded as Clinical Concept, Incorrectly Explained), and 2 157 Clinical Trial errors and 8 of 220 Visit Note errors were assigned two codes. When judgments required specialized clinical knowledge, such as in the case of clinical concepts, findings, and diagnoses, a clinician reviewed the coding. If the clinician suggested changes, these changes were made, which resulted in seven corrections in the researchers' original codes. Finally, to assess inter-rater reliability of the procedure, two coders independently coded 5 Clinical Trial and 5 Visit Notes protocols, which included 15 and 17 errors, respectively. Several months passed between the initial coding and the inter-rater reliability check, making it highly unlikely that the coders could simply recall their original codes. Based on Neuendorf [65] and Potter and Levine-Donnerstein [66]guidelines for assessing inter-rater agreement in content analysis, simple percent agreement was used as the inter-rater reliability measure. Agreement level was acceptable, 93% for Clinical Trial and 82% for Visit Notes. Disagreements were resolved through discussion.

4 RESULTS

Number of errors in each error subcategory and the number of participants who made those errors are presented in Table 3.

4.1 Examples of errors in each category

Qualitative analysis of the errors is presented below.

4.1.1 Errors in describing clinical concepts

4.1.1.1 Clinical Trial (35 errors, made by 25 participants, or 31% of all participants): Errors in this category involved incorrect explanations of some clinical concepts from the original texts. Many incorrect explanations concerned the nature of diabetes and the function of the liver, and the pertinence of liver functioning to diabetes (glycogen metabolism). For example, with respect to explaining the nature of diabetes, two participants stated that it was "a disorder of glucose production," rather than of glucose metabolism. Another participant wrote that "some diabetics have a liver disease," while another suggested that diabetes damages the liver. Several stated that the function of the liver is to produce either glucose or insulin, and that it is this function that is impaired in diabetes.

A related, and probably underlying, error involved explanation of various chemical substances involved in glucose metabolism. Several participants seemed to confuse glucose and glycogen with insulin and did not understand that glycogen was a form of glucose, rather than a separate substance involved in glucose metabolism. This is exemplified by the following statement from one of the participants, "A normal body will after a meal store extra *insulin* in the muscles and liver as glycogen." Another participant wrote that "glucose breaks down consumed calories." Yet another wrote that in diabetics "sugar drops after eating."

Other errors involved nuanced variations on the ones described above, such as stating that conversion of glucose to glycogen is done by "blood cells."

4.1.1.2 Visit Note (18 errors, made by 15 participants, or 19% of all participants): Errors in this category often had to do with the explanation of cardiovascular functioning and diabetes. Examples of errors in the explanation of cardiovascular functioning involve equating irregular heartbeat with hurt murmur, stating that the patient's heart rate "was irregular…, **especially in the lower area of his heart** [bold face ours, for emphasis]," or referring to "the apex of heart beat" and thus misrepresenting the location where heartbeat was heard best as its characteristic.

Misconceptions in explaining diabetes-related concepts were somewhat similar to those demonstrated in retelling the clinical trial. Examples include implying that diabetes was a liver disorder ("diabetes medications for his liver") and misunderstanding the role of insulin in diabetes, thus making errors when referring to the two diabetes medications prescribed in the visit ("something to help with absorption of insulin.")

Other errors involved misattributing causality of the processes described in the note. For example, one participant stated, "He had about 18 breaths per minute, for his respiration rate, which may explain his constant wheezing and also that he quickly becomes short of breath." The correctly cited respiration rate of 18 breaths per minute is actually within the normal range. In two cases, conceptual errors involved misattributing the primary cause of edema (swelling) in the legs to the patient's diabetes, pneumonia, and being overweight. While these may be contributing factors, the likeliest primary cause in this scenario is cardiac dysfunction.

A few misconceptions had to do with organs / symptoms / functions unrelated to the heart, breathing or diabetes (e.g., “The neck examination was abnormal, **probably caused by mucus build up**…” [Bold face ours, for emphasis]).

4.1.2 Errors in describing research purpose

4.1.2.1 Clinical Trial (42 errors, made by 41 participant, or 51% of all participants): This error type was unique to the clinical trial document, since thiswas the only document describing a clinical research study. Forty-two participants incorrectly explained the purpose of the trial. While the actual objective of the trial was to test an assessment procedure, 29 participants (36% of all) believed that this was a treatment trial. As stated by one of the participants, “This article is about a new technique to help patients with diabetes’ liver process sugar.” When describing the purpose of the trial, participants often used terms such as “treatment” and “medication.” Statements of five participants suggested that the trial is for diagnosing a type of diabetes or “this condition,” such as “The new assessment they wish to develop, in hope that it would be more affordable and easily available to diagnose this type of diabetes... “ One participant suggested that the trial tested a diagnostic procedure for liver disease. The remaining erroneous purpose statements varied. For example, some participants correctly stated that the trial tests an assessment procedure, but misrepresented exactly what was being assessed (e.g., “blood sugar”). Another suggested that the procedure under investigation was being performed to see “where the glucose is being stored.”

Most cases of errors in retelling the purpose of the trial were unambiguous. However, a few retellings involved seemingly contradictory statements; these participants stated in one sentence that that the objective of the trial was to develop an assessment method, and in another sentence, that the purpose was to develop a new treatment.

4.1.3 Errors in reporting findings and diagnoses

4.1.3.1 Visit Note (Findings -52 errors, made by 45 participants, or 61% of all participants; Diagnoses - 12 errors, made by 12 participant, or 15% of all participants): Because of the nature of the documents used in this study, all findings and diagnoses codes pertain to the VisitNote document only.

4.1.3.1.1 Finding, inaccurately reported, FIR (31 errors by 25 participants, or 31% of participants): Errors in this category involve clinically possible misrepresentation of a Note’s finding. The most common type of inaccurately reported finding involved misinterpreting the nature or misreporting the location of a symptom, or adding a descriptor or a circumstance to the original finding. Examples of misinterpreting the nature of a symptom involve writing “new II/IV diastolic murmur is observed”, when the observed murmur is systolic, or stating that there are “problems with systolic pressure,” when the note indicates “LV systolic dysfunction.” Cases of misreported locations are exemplified by a statement such as “his head is normal with normal carotid uptake.” Here, “carotid uptake” is an incorrect term for “carotid upstroke,” or pulsation in the carotid artery, which supplies oxygenated blood to the head and neck, and is measured in the neck (rather than the head). Another example of a misreported location is “one swollen foot” for “pedal edema,” or swollen feet – a symptom that typically affects both feet. The error was most likely caused by misinterpreting the “1+” shorthand descriptor of the pedal edema in the document, which refers to the severity of edema, rather than to the number of extremities that it affects. One example of adding incorrect descriptors or circumstances is the statement that the patient being described has “*upper* respiratory problems”. In fact, the Note states that the patient is recovering from pneumonia, which is a disease of the lungs rather than the upper respiratory pathways. Another example is the statement that the patient experiences

"difficulty breathing *when standing up*." According to the Note, the patient experiences orthopnea, or severe breathing difficulty *when lying down*.

Other common errors in this category involved interpreting and / or abnormal findings as normal, or normal as abnormal, or reversing the direction or changing a descriptor provided in a finding. Examples of abnormal-to-normal substitutions involve stating that the patient had healthy kidneys (the Note reports diabetic nephropathy and renal failure) and normal heart noises (the Note reports heart murmur). Normal-to-abnormal reversal is illustrated by statements that the patient had abnormal heart rate (reported in the Note as 75-85, which is normal) or high blood pressure (120/60 according to the Note, again, a value within normal range). Other types of reversals and descriptor changes are exemplified by statements that the patient had "something hard in the abdomen area" (the Note states: Abdomen: soft) and had "firm" neck (the Note stated that the neck was "supple").

4.1.3.1.2 Finding, Non-existent, FNE (9 errors by 8 participants, or 10% of participants): Nonexistent findings are clinically possible findings that represent a mere "partial" recall or a distortion of existing findings. In some cases, they cannot be traced to any specific statement in the Note. In other cases, it is possible to detect what information in the Note could have given rise to them, but they demonstrate an extreme inference or complete misunderstanding. Eight participants made nine instances of this error. Examples include stating that the patient had "at least one foot sore" (perhaps, as mentioned above, prompted by the Note statement of 1+ pedal edema) or "orthopedic troubles;" suffered from pain, had "abnormal nerve function" (a confusion of "nephropathy" with similarly sounding "neuropathy"?) or needed to urinate frequently (a finding which is potentially possible, given the diagnosis of Diabetes Mellitus, but not actually stated in the Visit Note).

4.1.3.1.3 Finding, non-sensical, FNS (12 errors by 12 participants, or 15% of participants): Errors in this category are clinically impossible, in a way that is apparent to many non-health professionals. Examples include stating that there was "the sound of bones scraping together" in the lungs, the patient "was diagnosed with chronic Coumadin", "had trouble breathing a green expectorant" or had "flat jugular vein".

4.1.3.1.4 Diagnosis, inaccurately reported, DiaIR (7 errors by 7 participants, or 9% of participants): This coding category involves misrepresentation of some characteristics of a diagnosis, mentioned in the Note. The code is similar to "Finding, inaccurately reported", but is a misrepresentation of an explicit diagnosis, rather than a finding. Seven participants made seven errors in this category. Six of seven inaccuracies occurred in reporting the cardiovascular diagnoses, which in the text were stated as "atherosclerotic coronary vascular disease with old myocardial infarction" and "ischemic heart disease." The inaccurate reports ranged from "atherovascular… with past myocardial infarction," to "his heart seemed to be OK." One non-cardiac inaccuracy was the addition of a modifier to "renal failure," thus turning it into "severe kidney failure."

4.1.3.1.5 Diagnosis, wrong, DiaW (4 errors by 4 participants, or 5% of participants): This error category involved attributing as a diagnosis a disease or a condition that was not mentioned in the original document. All instances involved stating that the patient suffered from neuropathy or diabetic (one specifying "diabetic neuropathy in his feet"). Like the errors mentioned above, these mistakes were most probably the result of terminology confusion between similarly sounding "diabetic nephropathy" and "diabetic neuropathy." We chose to place these into the wrong diagnosis category rather than the terminology confusion category because this error resulted in the participant's identifying a diagnosis that really existed, but was not given in the Note.

4.1.3.1.6 Diagnosis, non-sensical, DiaNS (1 errors by 1 participants, or1% of participants): This category involved clinically impossible / meaningless diagnoses. The one instance in this category was "artherosclerotic disease with terminal PL ischemia."

4.1.4 Errors in Medication Names—Jointly across sub-codes, errors in medication names comprised the largest single error category, which included 93 errors. For the Clinical Trial document, there were 24 errors, made by 24 participants, or 30% of the participants. For the Visit Note, there were 69 errors, made by 57 participants, or 71% of the participants. Errors belonging to three of the subcategories, Medication, non-prescribed (MedNP), Medication generic, misspelling (MedGMiss) , and Medication nature, incorrect (MedNI) were found in retellings of both document types. The remaining error categories, Medication, brand name confusion (MedBrC), Medication, brand name confusion (MedBrC), Medication, brand name misspelling (MedBrM), Medication, partial memory (MedPM) were in the Visit Notes' retelling only. This was the case because these categories all pertained to errors in medications' brand names, and the Clinical Trial document did not make references to medication brands.

4.1.4.1 Medication, non-prescribed, MedNP

4.1.4.1.1 Clinical Trial (3 errors by 3 participants, or 4% of participants): All errors in this category involved substituting one generic medication name (acetaminophen) with another (aspirin)

4.1.4.1.2 Visit Note (14 errors by 12 participants, or 15% of participants): The most common non-prescribed medication error involved referring to a medication's purpose, when the purpose did not correspond to anything actually described in the Note. Examples include "medicine for [the patient's] kidneys," "medication for liver functioning problems," "medication to prevent decay of the vessels," etc. Typically, though not always, the named medications were related to a condition that the patient, indeed, was reported to be suffering from, such as heart, lung, and kidney problems. There were nine errors of this type.

There were also a few instances of mentioning a non-prescribed medication type ("an antibiotic"), a generic name ("acetometaphin," participant's spelling), or form ("some nasal spray"). The source of these errors is not easily traceable to anything in the Note.

4.1.4.2 Medication generic, misspelling, MedGMiss

4.1.4.2.1 Clinical Trial (19 errors by 19 participants, or 24% of participants): Acetaminophen, the one generic medication name in the Clinical Trial document, produced all instances of this error. Noteworthy, each error was unique, resulting in the nineteen different misspellings of "acetaminophen."

4.1.4.2.2 Visit Note (1 error by 1 participant, or 1% of participants): The only error in this category in the VisitNote retelling involved misspelling "acetaminophen" as "acetometaphin."

Box A presents the total of 20 ways to misspell "acetaminophen" demonstrated in this study.

4.1.4.3 Medication nature, incorrect, MedNI

4.1.4.3.1 Clinical Trial (2 errors by 2 participants, or 3% of participants): In the case of the clinical trial document retelling, this error type involved a confusion between water (H2O) and heavy, or hydrogen-enriched, water (H2O2), as well as calling heavy water "a medicine."

4.1.4.3.2 Visit Note (1 errors by 1 participant, or 1% of participants): This category involved ascribing an incorrect function to a prescribed medication. The one instance of this error type involved describing Coumadin as “sleep medication.”

4.1.4.4 Medication, brand name confusion, MedBrC

4.1.4.4.1 Visit Note (9 errors by 8 participant, or 10% of participants): Errors in this category involve “swapping” two existing brand names with each other, or mentioning a brand that exists, but was not mentioned in the Note. The Note mentions prescribing Flovent, a medication to treat lung inflammation. The most common error involved brand names that sounded similar to Flovent, including Flomax (commonly prescribed to treat the symptoms of an enlarged prostate, such as difficulty urinating), Flonase (a corticosteroid for treating allergy symptoms), and Flovin (a ciprofloxacin, an antibiotic brand not available in the US). Another example of brand name confusion involves the statement that the patient was prescribed Dilantin (possibly, a combination of Diovan and Lantus, present in the Note).

4.1.4.5 Medication, brand name misspelling, MedBrM

4.1.4.5.1 Visit Note (18 errors by 17 participant, or 21% of participants): This category involves spelling errors in brand names that can be corrected by replacing or deleting one letter. Examples include misspelling Coumadin as “Cumodin” or “Comadin”, or Ambien as “Ambient.”

4.1.4.6 Medication, partial memory, MedPM

4.1.4.6.1 Visit Note (24 errors by 16 participant, or 20% of participants): This category involves errors in medication brand names, which render the source name recognizable if one is familiar with the Note, but which are more extensive errors than replacing or deleting one letter. Examples include “Novil” for “Novolin”, “Courdin” for “Coumadin,” and “Torval” for “Torpol.”

4.1.5 Medical terminology

4.1.5.1 Clinical Trial (33 errors by 27 participant, or 34% of participants): This error category involved substituting a specialized medical term with another medical or general term, or making an error in an abbreviation. Common instances included confusing hypoglycemia and hyperglycemia (stating that “too much sugar in the blood” is called “hypoglycemia”); confusing glucose and glycogen, or glycogen and glycerin, and thus, for example, referring to glycogen as “sugar in the blood”; and calling the spectroscopy procedure “spectronomy,” “stenoscopy,” or “spectral chromatography.” Other errors involved providing an incorrect chemical formula for heavy water (e.g., “HC2O”), or giving an incorrect abbreviation for the 13C-MRS procedure.

4.1.5.2 Visit Note (16 errors by 15 participant, or 19% of participants): Incorrectly applied terms range from stating that the patient was “coughing with green expectorant”, to calling “Diabetes Mellitus” “mettice” and “myocardial infarction” “myopathic infarction.” Many errors also involved an incorrect abbreviation for “ICD” implant. One participant confused medication name “Lasix” with “Lasik”, a surgical procedure; another wrote that “Apneia” was a medication name.

5 DISCUSSION AND CONCLUSIONS

The objective of this study was to develop a data-driven classification scheme for lay comprehension errors, derived on the basis of two representative medical documents,

intended for patients' reading under participatory health care model. Without being a comprehensive taxonomy, this scheme provides a starting point for the important task of categorizing and remediating such errors. From a practical perspective, this work provides a description and an insight into possible causes of several error types than can be remediated via a combination of educational and informatics approaches. From a theoretical perspective, we provide a proof of concept for methodological feasibility of the task. Future efforts at developing a comprehensive taxonomy may use the combination of document retelling and content analysis, but apply it to a broader range of document types and a greater number of documents. As our documents had little numerical data, we were not able to elucidate common numeracy-related errors, such as problems with interpreting graphs and risk values and dosage conversions. Participants in our study had above-average education level, were healthy, and did not have an intrinsic motivation to understand the two documents. We should also be cautious interpreting the results of a study where all participants were recruited at a single academic site. At the same time, because it included not only students and faculty, but also staff, our sample was broader than a typical university sample. Their healthy volunteers' status and thus insufficient motivation could increase the number of comprehension errors, while the level of education had the potential of decreasing it. Future studies should draw upon a more diverse sample of participants, which would include more variability of demographic characteristics and health status.

Besides the narrow scope, limitations of our study involve some overlap between categories. For example, confusion between two similarly sounding terms, such as hypoglycemia and hyperglycemia--most likely caused by the lack of familiarity with professional medical terminology--was categorized as a terminology error. Statements that demonstrated lack of understanding of scientific concepts (e.g., functions of hormones and the role of insulin in sugar metabolism) were classified as errors in clinical concepts. Although incorrect usage of the terms or incorrect conceptual explanations often *constitute* errors, they also often *underlie* or *cause* errors in statements about findings, diagnoses, and procedures. When a participant says that the patient had been diagnosed with "diabetic neuropathy," this confusion of "nephropathy" and "neuropathy", classified as an error in diagnosis, is likely to be caused by the lack of terminological knowledge. Similarly, insufficient biological knowledge and non-normative beliefs (or theories) of health and disease not only clearly underlie errors in clinical concepts, but may also be at the root of findings and diagnosis-related errors. This blurring of categorical borders in our scheme is related to the difficulty in distinguishing between the two aspects of errors, as described by JCAHO [40], 1) things that go wrong and 2) reasons why this happens. The difficulty is methodological and relates to the nature of the task, that is, reading comprehension. In comprehension, causes of errors are largely cognitive, related to lapses in memory and attention and insufficient knowledge. However, the outcomes are also cognitive (e.g., accuracy of verbal answers). As a result, distinguishing between the cause (specific health beliefs and theories or lack of terminology knowledge) and the effect of a misunderstanding (e.g., incorrectly recalled diagnosis or misuse of a medical term in an explanation of a disease mechanism) is more difficult than in medical errors that involve actions (e.g., administering a wrong medication with a similarly-sounding name, due to terminology confusion). This also results in a scheme where categories differ in their "depth", as some error classifications are likely to be related to lay beliefs and theories, while others (e.g., misspellings) may be more straightforward in their origins. A two-dimensional taxonomy, distinguishing between the causes of errors and the errors themselves, would address the issue of overlap, but developing it is challenging for the same reasons that cause the existing overlap.

The existing classification scheme has some insights for development of electronic medical documents. While PHRs and other online consumer health resources have the potential of improving patients' and consumers' experience with participatory healthcare, findings of

this study suggest that lay people need support comprehending medical documents and, by extrapolation, authoring documents. Even in this group of highly educated participants, comprehension errors were frequent. They were also broad in scope, including understanding of research conventions, biomedical concepts, medical facts, and professional medical terms. Due to the largely narrative nature of our documents, we did not record many numeracy-related errors, but work by other authors suggests that problems with numeric conversions (e.g., dosages) and data representations are also common ([43]). The diverse range of error types suggests that informatics support to document comprehension should be multi-faceted, and at the same time tailored to specific problems. At the present time, most tools are directed at translating professional medical terminology into consumer-friendly terms (e.g., Zeng-Treitler and colleagues [27]). This appears to be necessary, but not sufficient: EHR/PHR and informed consent documents also need contextually relevant educational materials, easy to read summaries of findings and their interpretations, explanation of ranges and values of tests results, glossaries of medications' names, and more.

While working on improving laypeoples' experience with medical documents, it is also important for informaticians and educators to remember that making patients' understanding of the documents mirror that of their healthcare providers is neither realistic, nor desirable. In the absence of specialized biomedical knowledge and clinical experience, laypeople will ascribe different level of importance to different statements, remember different facts, and organize information differently. The goal of lay comprehension support is not to position patients as professionals, but to enable them to work with professionals in the most effective way possible. We should also keep in mind that differences in lay and professional views on treatment and care should not be reduced to patient comprehension errors. Patients and healthcare professionals have different models of health and disease, which may produce somewhat discrepant value and belief systems. Physicians' reasoning and decision-making is guided by the "disease model," in which health problems are prominently connected to pathophysiological mechanisms. Patients, on the other hand, are guided by the "illness model," which may include a combination of formal knowledge, naïve health concepts ("folk biology"), personal experience, and social and emotional implications of the disruption of normal routines caused by the illness (Patel, Arocha, & Kushniruk [68]). While discussing patient-provider communication across these different belief systems is beyond the scope of this work, it is important to be aware that not all misunderstandings can be corrected via informatics support of "plain language" in medical documents.

This study classified participants' comprehension errors into nine categories and twenty three subcategories; there are other possible error types that were not elucidated here because of the nature of the two documents. Not all error types are equally critical for comprehension and not all are likely to have similar impact on health behavior and decision making. Of the errors in this study, the one raising the most concern is interpreting the aim of the clinical trial. Although the purpose of the trial was phrased in the document as "to establish a new assessment method for glycogen metabolism," it was frequently misconstrued as developing treatment. Twenty nine of the eighty participants made that error; another twelve participants misinterpreted the purpose of the trial without defining it as treatment. Participants in our study were healthy, well-educated individuals, who were reading the description of the trial in a comfortable emotional state, without the overwhelming anxiety that would accompany a bad diagnosis given to themselves or to a loved one. If these people had difficulty understanding the trial's objectives, we should expect similar or greater difficulty in the general population of patients and caregivers. Misunderstanding objectives of clinical trials and routine procedures is likely to lead people to enroll into trials without true informed consent or seek inappropriate treatment. It may also lead to a targeted population missing information about trials of potential interest: if lay

people do not understand the research purpose of trial descriptions, they are not likely to bring those trials to their physicians' attention and ask whether the trials are right for them. While pre-empting this error is critical, it is also challenging without careful one-on-one involvement with a health professional. Attempts to explore informatics solutions may rely on cognitive science research of the impact of text signals (e.g., bullets, section headings, color highlights) on comprehension and attempt to use these signals to make research objectives more prominent. The problem of misunderstanding research objectives, however, cannot be corrected by information alone, as it is likely to be tied to the public's lack of understanding of clinical research objectives and conventions. This study suggests the importance of discussing the rationale behind clinical research, benefits to society vs. direct benefits to participants, and different types of research studies (e.g., interventional vs. observational) in health and science education.

Another error category, common in this study, involved medication-related errors. Twenty-four of eighty participants made medication errors when reading the clinical trial document; fifty-seven of eighty made them when reading the visit note. Most errors had to do with medication names, particularly with misspelling and confusing brand names of medications. One of the limitations of this study is that it is impossible to ascertain whether the misspellings were, indeed, knowledge errors (participants did not *know* how to spell the name) or attentional slips resulting in typographic errors. While we are not aware of research into consumers' confusion of medication names, as opposed to confusion in pharmacies [69] and medical centers, the extreme human impact of medication errors ensures that the phenomenon is well-studied in the medical literature that focuses on health professionals. For example, Koczmara and Hyland [70] report on confusion of two particular drug names – Plavix and Pradax – and the implications for critical care nursing. Senger et al.[71] write about drug misspellings as an information retrieval problem in Heidelberg University's drug information system, describing error types as cognitive, phonetic and typographic, with typographic the most problematic. "Look-alike, sound-alike" (LASA) errors are defined by Basco et al. [72] as "the erroneous prescription or delivery of a drug because the name of the drug (generic or brand) is similar in appearance to or sounds like another drug". Not surprisingly, lay people have difficulty with aspects of medication names that are challenging to health professionals, and need help distinguishing among similarly-sounding medications and dealing with spelling of medication names. In most consumer interactions with medication names lay individuals have to state medication names orally or recognize their spelling by professionals, rather than spelling the names themselves. However, commonality of look-alike, sound-alike medication names suggests that the issue requires further investigation. It is tempting to speculate about the role played by DTC (direct to consumer) marketing of pharmaceuticals in the behavior of our participants. Regardless, these findings have implications for consumers' information retrieval behavior and ability to self-educate using materials they find online. Consumers who have confused the Flovent they really need with the Flomax they half-remember will be unable to make much sense of information about either medication in any medium. Compared with understanding of research objectives and clinical concepts, supporting recall of medication names is within an easier reach for informatics. PHRs can include medication name spell-checkers, specify medication function, and provide names of similarly sounding medications and ask verification questions.

In discussing the theoretical issues around our classification scheme, we mentioned some overlap between two aspects of errors: the "what" and the "why" of things going wrong, or the underlying causes and the stated inaccuracies themselves. In our classification scheme statements that are clearly indicative of misunderstanding clinical concepts and not pertaining to findings, diagnoses, devices or procedures, make up a separate coding category. However, insufficient conceptual / biological knowledge or non-normative beliefs

(theories) of health and disease are also likely to be the reason behind many errors related to misreporting findings and diagnoses. Knowledge of biological concepts is typically acquired over years of formal education, and is best addressed in the K-12 educational system. Research in science education suggests that solid biological knowledge, indeed, often underlies accurate health reasoning and effective health information seeking [73, 74] However, when point-of-care remediation is necessary, informatics can provide it via tailored and contextualized educational materials, exemplified by Baorto, Li and Cimino's [26].

In summary, this study suggests that lay people have difficulty reading medical documents, comprehension of which is essential for meaningful participation in their care. It also suggest that errors that people make can be classified into a manageable number of hierarchical categories, which are useful for thinking about ways to support lay users of electronic medical documents. Most common errors made by the participants in this study pertained to understanding conventions and objectives of clinical research, knowledge of health concepts and corresponding recall of medical findings and diagnosis, medical terminology and spelling, and problems with medication names. Future research will fine-tune these categories and identify new challenging areas, supporting tools for helping patients and consumers dealing medical documents.

Acknowledgments

The authors gratefully acknowledge Marsha Naydich, MD for her assistance with the data coding. Funding support by NIDDK 7R01DK075837; and the Intramural Research Program of the National Library of Medicine.

References

1. U.S. Department of Health and Human Services, National Cancer Institute, and National Committee on Vital and Health Statistics. Personal health records and personal health record systems. Bethesda, MD: U.S Department of Health and Human Services; 2006 Feb. Available online: http://www.himss.org/asp/contentredirector.asp?contentid=78505
2. Brennan PF, Downs S, Casper G, Kenron D. Project HealthDesign: stimulating the next generation of personal health records. AMIA Annu Symp Proc. 2007 Oct 11.:70–4. [PubMed: 18693800]
3. Brennan PF, Casper G, Downs S, Aulahk V. Project HealthDesign: enhancing action through information. Stud Health Technol Inform. 2009; 146:214–8. [PubMed: 19592837]
4. Hampton T. Groups push physicians and patients to embrace electronic health records. JAMA. 2008 Feb 6; 299(5):507–9. [PubMed: 18252875]
5. Stevens MM. "Shuttle sheet": a patient-held medical record for pediatric oncology families. Med Pediatr Oncol. 1992; 20(4):330–5. [PubMed: 1608356]
6. Kaelber DC, Jha AK, Johnston D, Middleton B, Bates DW. A research agenda for personal health records (PHRs). J Am Med Inform Assoc. 2008 Nov-Dec;15(6):729–36. [PubMed: 18756002]
7. Health Information Management Systems Society. [November 6, 2011] HIMSS Personal Health Records Definition and Position Statement. 2007 Sep 28. Available online: http://www.himss.org/content/files/PHRDefinition071707.pdf
8. Marshall P. Keeping tabs. How personal health records are changing the face of healthcare. Mark Health Serv. 2009 Winter;29(4):20–5. [PubMed: 20043413]
9. Wynia MK, Torres GW, Lemieux J. Many physicians are willing to use patients' electronic personal health records, but doctors differ by location, gender, and practice. Health Aff (Millwood). 2011 Feb; 30(2):266–73. Erratum in: Health Aff (Millwood). 2011 Jun;30(6):1217. [PubMed: 21289348]
10. Wuerdeman L, Volk L, Pizziferri L, Tsurikova R, Harris C, Feygin R, Epstein M, Meyers K, Wald JS, Lansky D, Bates DW. How accurate is information that patients contribute to their Electronic Health Record? AMIA Annu Symp Proc. 2005:834–8. [PubMed: 16779157]
11. Kim MI, Johnson KB. Patient entry of information: evaluation of user interfaces. J Med Internet Res. 2004 May 14.6(2):e13. [PubMed: 15249262]

12. Kim EH, Stolyar A, Lober WB, Herbaugh AL, Shinstrom SE, Zierler BK, Soh CB, Kim Y. Challenges to using an electronic personal health record by a low-income elderly population. J Med Internet Res. 2009 Oct 27.11(4):e44. [PubMed: 19861298]
13. Tran DT, Zhang X, Stolyar A, Lober WB. Patient-centered design for a personal health record system. AMIA Annu Symp Proc. 2005:1140. [PubMed: 16779426]
14. Lober WB, Zierler B, Herbaugh A, Shinstrom SE, Stolyar A, Kim EH, Kim Y. Barriers to the use of a personal health record by an elderly population. AMIA Annu Symp Proc. 2006:514–8. [PubMed: 17238394]
15. Britto MT, Jimison HB, Munafo JK, Wissman J, Rogers ML, Hersh W. Usability testing finds problems for novice users of pediatric portals. J Am Med Inform Assoc. 2009 Sep-Oct;16(5):660–9. Epub 2009 Jun 30. [PubMed: 19567793]
16. Ming ME, Levy RM, Hoffstad OJ, Filip J, Gimotty PA, Margolis DJ. Validity of patient self-reported history of skin cancer. Arch Dermatol. 2004 Jun; 140(6):730–5. [PubMed: 15210466]
17. Khoja S, McGregor SE, Hilsden RJ. Validation of self-reported history of colorectal cancer screening. Can Fam Physician. 2007 Jul; 53(7):1192–7. [PubMed: 17872816]
18. Porter SC, Fleisher GR, Kohane IS, Mandl KD. The value of parental report for diagnosis and management of dehydration in the emergency department. Ann Emerg Med. 2003 Feb; 41(2):196–205. [PubMed: 12548269]
19. Ross SE, Lin CT. The effects of promoting patient access to medical records: a review. J Am Med Inform Assoc. 2003 Mar-Apr;10(2):129–38. [PubMed: 12595402]
20. Businger A, Buckel L, Gandhi T, Grant R, Poon E, Schnipper J, Volk LA, Wald JS, Middleton B. Patient review of selected electronic health record data improves visit experience. AMIA Annu Symp Proc. 2007 Oct 11.887
21. Keselman A, Slaughter L, Smith CA, Kim H, Divita G, Browne A, Tsai C, Zeng-Treitler Q. Towards consumer-friendly PHRs: patients' experience with reviewing their health records. AMIA Annu Symp Proc. 2007 Oct 11.:399–403. [PubMed: 18693866]
22. Archer N, Fevrier-Thomas U, Lokker C, McKibbon KA, Straus SE. Personal health records: a scoping review. J Am Med Inform Assoc. 2011; 18:515–522. [PubMed: 21672914]
23. Lee M, Delaney C, Moorhead S. Building a personal health record from a nursing perspective. Int J Med Inform. 2007 Oct; 76(Suppl 2):S308–16. Epub 2007 Jul 5. [PubMed: 17616432]
24. Marchionini, G.; Rimer, BK.; Wildemuth, B. [November 6, 2011] Evidence base for personal health record usability: Final Report to the National Cancer Institute. 2007. Available online: http://ils.unc.edu/phr/files/final%20report%20010307.pdf
25. Maloney FL, Wright A. USB-based Personal Health Records: an analysis of features and functionality. Int J Med Inform. 2010 Feb; 79(2):97–111. Epub 2010 Jan 6. [PubMed: 20053582]
26. Baorto D, Li L, Cimino JJ. Practical experience with the maintenance and auditing of a large medical ontology. J Biomed Inform. 2009 Jun; 42(3):494–503. Epub 2009 Mar 12. [PubMed: 19285569]
27. Zeng-Treitler Q, Goryachev S, Kim H, Keselman A, Rosendale D. Making texts in electronic health records comprehensible to consumers: a prototype translator. AMIA Annu Symp Proc. 2007 Oct 11.:846–50. [PubMed: 18693956]
28. Cardillo E, Tamilin A, Serafini L. A methodology for knowledge acquisition in consumer-oriented healthcare. Communications in Computer and Information Science. 2011; 128(Part 3):249–261.
29. Brennan PF, Downs S, Casper G. Project Health Design: rethinking the power and potential of personal health records. J Biomed Inform. 2010 Oct; 43(5 Suppl):S3–5. [PubMed: 20937482]
30. Adnan M, Warren J, Orr M. Enhancing patient readability of discharge summaries with automatically generated hyperlinks. Health Care and Informatics Review Online. 2009; 13(4)
31. Paasche-Orlow MK, Taylor HA, Brancati FL. Readability standards for informed-consent forms as compared with actual readability. N Engl J Med. 2003 Feb 20; 348(8):721–6. [PubMed: 12594317]
32. Joffe S, Cook EF, Cleary PD, Clark JW, Weeks JC. Quality of informed consent in cancer clinical trials: a cross-sectional survey. Lancet. 2001 Nov 24; 358(9295):1772–7. [PubMed: 11734235]
33. Olver IN, Whitford HS, Denson LA, Peterson MJ, Olver SI. Improving informed consent to chemotherapy: a randomized controlled trial of written information versus an interactive

multimedia CD-ROM. Patient Educ Couns. 2009 Feb; 74(2):197–204. Epub 2008 Oct 21. [PubMed: 18945572]

34. Flory J, Emanuel E. Interventions to improve research participants' understanding in informed consent for research: a systematic review. JAMA. 2004 Oct 6; 292(13):1593–601. [PubMed: 15467062]
35. Kang EY, Fields HW, Kiyak A, Beck FM, Firestone AR. Informed consent recall and comprehension in orthodontics: traditional vs improved readability and processability methods. Am J Orthod Dentofacial Orthop. 2009 Oct; 136(4):488.e1–13. discussion 488-9. [PubMed: 19815144]
36. Dovey SM, Meyers DS, Phillips RL Jr, Green LA, Fryer GE, Galliher JM, Kappus J, Grob P. A preliminary taxonomy of medical errors in family practice. Qual Saf Health Care. 2002; 11:233–238. [PubMed: 12486987]
37. Benner P, Sheets V, Uris P, Malloch K, Schwed K, Jamison D. Individual, practice, and system causes of errors in nursing. A taxonomy. J Nurs Adm. 2002; 32:509–523. [PubMed: 12394596]
38. Ely JW, Levinson W, Elder NC, Mainous AG III, Vinson DC. Perceived causes of family physicians' errors. J Fam Pract. 1995; 40:337–344. [PubMed: 7699346]
39. Davis T. Paediatric prescribing errors. Arch Dis Child. 2011; 96:489–491. [PubMed: 21109509]
40. Chang A, Schyve PM, Croteau RJ, O'Leary DS, Loeb JM. The JCAHO patient safety event taxonomy: a standardized terminology and classification schema for near misses and adverse events. Int J Qual Health Care. 2005 Apr; 17(2):93–4. [PubMed: 15772256]
41. Zhang J, Patel VL, Johnson TR, Shortliffe EH. A cognitive taxonomy of medical errors. J Biomed Inform. 2004 Jun; 37(3):193–204. [PubMed: 15196483]
42. Chan CV, Matthews LA, Kaufman DR. A taxonomy characterizing complexity of consumer eHealth Literacy. AMIA Annu Symp Proc. 2009 Nov 14.2009:86–90. [PubMed: 20351828]
43. Ancker JS, Kaufman D. Rethinking health numeracy: a multidisciplinary literature review. J Am Med Inform Assoc. 2007 Nov-Dec;14(6):713–21. Epub 2007 Aug 21. [PubMed: 17712082]
44. Calhoun, Craig, editor. Dictionary of the Social Sciences. Cary, NC: Oxford University Press USA; 2002. p. 584
45. Tang PC, Black W, Young CY. Proposed criteria for reimbursing eVisits: content analysis of secure patient messages in a personal health record system. AMIA Annu Symp Proc. 2006:764–8. [PubMed: 17238444]
46. Lovasz N, Clarke J. Life beyond depression: the experience of gays and lesbians who self-identify as depressed. J LGBT Health Res. 2007; 3(4):53–73. [PubMed: 19042910]
47. Dukes RL, Borega KN, Clayton SL, Dukes CK, Fox C, McCombs J, Miller TL, Rodgers SE. Tainted love: content analysis of documents written by accused spousal abusers to their victims. Psychol Rep. 2002 Apr; 90(2):679–84. [PubMed: 12061614]
48. McGinn CA, Grenier S, Duplantie J, Shaw N, Sicotte C, Mathieu L, Leduc Y, Légaré F, Gagnon MP. Comparison of user groups' perspectives of barriers and facilitators to implementing electronic health records: a systematic review. BMC Med. 2011 Apr 28.9:46. [PubMed: 21524315]
49. Appiah-Poku J, Newton S, Kass N. Participants' perceptions of research benefits in an african genetic epidemiology study. Dev World Bioeth. 2011 Dec; 11(3):128–35. [PubMed: 22103636]
50. Falk G, Brynhildsen J, Ivarsson AB. Contraceptive counselling to teenagers at abortion visits--a qualitative content analysis. Eur J Contracept Reprod Health Care. 2009 Oct; 14(5):357–64. [PubMed: 19916762]
51. Gebru K, Ahsberg E, Willman A. Nursing and medical documentation on patients' cultural background. J Clin Nurs. 2007 Nov; 16(11):2056–65. Epub 2007 Mar 1. [PubMed: 17331097]
52. Clarke EB, Luce JM, Curtis JR, Danis M, Levy M, Nelson J, Solomon MZ. Robert Wood Johnson Foundation Critical Care End-of-Life Peer Workgroup. A content analysis of forms, guidelines, and other materials documenting end-of-life care in intensive care units. J Crit Care. 2004 Jun; 19(2):108–17. [PubMed: 15236144]
53. Haas J, Bakken S, Bright TJ, Melton GB, Stetson P, Johnson SB. Clinicians' perceptions of usability of eNote. AMIA Annu Symp Proc. 2005; 973

54. Arar N, Seo J, Abboud HE, Parchman M, Noel P. Veterans' experience in using the online Surgeon General's family health history tool. Per Med. 2011 Sep 1; 8(5):523–532. [PubMed: 22076122]
55. Ayana M, Pound P, Ebrahim S. The views of therapists on the use of a patient-held record in the care of stroke patients. Clin Rehabil. 1998 Aug; 12(4):328–37. [PubMed: 9744668]
56. Wibe T, Hellesø R, Slaughter L, Ekstedt M. Lay people's experiences with reading their medical record. Soc Sci Med. 2011 May; 72(9):1570–3. Epub 2011 Mar 24. [PubMed: 21497971]
57. Bhavnani V, Fisher B, Winfield M, Seed P. How patients use access to their electronic GP record--a quantitative study. Fam Pract. 2011 Apr; 28(2):188–94. Epub 2010 Nov 17. [PubMed: 21084568]
58. Rassin M, Zilcha L, Berger M, Silner D. Personal medical documents management--how patients perceive, keep and manage their medical documents: a qualitative study. Int J Nurs Stud. 2007 Aug; 44(6):862–8. Epub 2006 May 5. [PubMed: 16678180]
59. Saboor S, Ammenwerth E. Developing a model for the adequate description of electronic communication in hospitals. Stud Health Technol Inform. 2011; 169:522–6. [PubMed: 21893804]
60. Whitson HE, Steinhauser K, Ammarell N, Whitaker D, Cousins SW, Ansah D, Sanders LL, Cohen HJ. Categorizing the effect of comorbidity: a qualitative study of individuals' experiences in a low-vision rehabilitation program. J Am Geriatr Soc. 2011 Oct; 59(10):1802–9. Epub 2011 Sep 13. 10.1111/j.1532-5415.2011.03602.x. [PubMed: 22091493]
61. Smith CA, Hetzel S, Dalrymple P, Keselman A. Beyond readability: investigating coherence of clinical text for consumers. J Med Internet Res. 2011 Dec 2.13(4):e104. [PubMed: 22138127]
62. Do NV, Barnhill R, Heermann-Do KA, Salzman KL, Gimbel RW. The military health system's personal health record pilot with Microsoft HealthVault and Google Health. J Am Med Inform Assoc. 2011 Mar-Apr;18(2):118–24. Epub 2011 Feb 2. [PubMed: 21292705]
63. Smith, CA. Doctoral dissertation. University of Pittsburgh: 2002. The clinical document architecture: XML semantic markup for enhanced clinical information retrieval.
64. Strauss, A.; Corbin, J. Basics of Qualitative Research: Grounded Theory Procedures and Techniques. Thousand Oaks, CA: Sage Publications; 1990. p. 288
65. Neuendorf, KA. The Content Analysis Guidebook. Thousand Oaks, CA: Sage Publications; 2002. p. 320
66. Potter JW, Deborah Levine-Donnerstein D. Rethinking validity and reliability in content analysis. Journal of Applied Communication Research. 1999; 273:258–284.
67. Lowe HJ, Antipov I, Hersh W, Smith CA, Mailhot M. Automated semantic indexing of imaging reports to support retrieval of medical images in the multimedia electronic medical record. Methods of Information in Medicine. 1999; 8:303–307. [PubMed: 10805018]
68. Patel VL, Arocha JF, Kushniruk AW. Patients' and physicians' understanding of health and biomedical concepts: relationship to the design of EMR systems. J Biomed Inform. 2002 Feb; 35(1):8–16. [PubMed: 12415722]
69. Lambert BL, Chang KY, Gupta P. Effects of frequency and similarity neighborhoods on pharmacists' visual perception of drug names. Soc Sci Med. 2003 Nov; 57(10):1939–55. [PubMed: 14499517]
70. Koczmara C, Hyland S. Drug name alert: potential for confusion between Pradax and plavix. Dynamics [serial online]. Sep; 2011 22(3):25–26. Available from: CINAHL Plus with Full Text, Ipswich, MA.
71. Senger C, Kaltschmidt J, Schmitt SP, Pruszydlo MG, Haefeli WE. Misspellings in drug information system queries: Characteristics of drug name spelling errors and strategies for their prevention. Int J Med Inform. 2010 Dec; 79(12):832–9. Epub 2010 Oct 15. [PubMed: 20951634]
72. Basco WT Jr, Ebeling M, Hulsey TC, Simpson K. Using pharmacy data to screen for look-alike, sound-alike substitution errors in pediatric prescriptions. Acad Pediatr. 2010 Jul-Aug;10(4):233–7. Epub 2010 Jun 11. [PubMed: 20541485]
73. Keselman A, Kaufman DR, Patel VL. "You can exercise your way out of HIV" and other stories: The role of biological knowledge in adolescents' evaluation of myths. Sci Ed. 2004; 88(4):548–573.

74. Keselman A, Browne AC, Kaufman DR. Consumer health information seeking as hypothesis testing. J Am Med Inform Assoc. 2008 Jul-Aug;15(4):484–95. Epub 2008 Apr 24. [PubMed: 18436912]

Textbox 1

Clinical Trial document

NCT00481598 Non Invasive Assessment of Liver Glycogen Kinetics in Type1 Diabetics

Patients with Type 1 diabetes suffer from impaired postprandial hepatic glycogen storage and breakdown, if they are under poor glycaemic control. Poor glycogen storage in the liver puts these patients at risk of fasting hypoglycemia. Amelioration of glycaemic control could improve these abnormalities and thereby reduce the risk of hypoglycemia in these patients. The “gold standard” technique for the assessment of hepatic glycogen metabolism in humans, 13 C magnetic resonance spectroscopy (13C-MRS), is expensive and limited to a few centers worldwide. Aim 1 of our project is to establish a new assessment method for glycogen metabolism. This new method is based on oral administration of 2H2O and acetaminophen.

Textbox 2

Visit Note document

History of Present Illness

This 66-year-old white male was seen in my office on Month DD, YYYY. Patient was recently discharged from Doctors Hospital at Parkway after he was treated for pneumonia. Patient continues to have severe orthopnea, paroxysmal nocturnal dyspnea, cough with greenish expectoration. His exercise tolerance is about two to three yards for shortness of breath. The patient stopped taking Coumadin for reasons not very clear to him. He was documented to have recent atrial fibrillation. Patient has longstanding history of ischemic heart disease, end-stage LV systolic dysfunction, and is status post ICD implantation. Fasting blood sugar this morning is 130.

Physical Examination

VITAL SIGNS: Blood pressure is 120/60. Respirations 18 per minute. Heart rate 75-85 beats per minute, irregular. Weight 207 pounds.

HEENT: Head normocephalic. Eyes, no evidence of anemia or jaundice. Oral hygiene is good.

NECK: Supple. JVP is flat. Carotid upstroke is good.

LUNGS: Severe inspiratory and expiratory wheezing heard throughout the lung fields. Fine crepitations heard at the base of the lungs on both sides.

CARDIOVASCULAR: PMI felt in fifth left intercostal space 0.5-inch lateral to midclavicular line. First and second heart sounds are normal in character. There is a II/VI systolic murmur best heard at the apex.

ABDOMEN: Soft. There is no hepatosplenomegaly.

EXTREMITIES: Patient has 1+ pedal edema.

Medications

1. Ambien 10 mg at bedtime p.r.n.
2. Coumadin 7.5 mg daily.
3. Diovan 320 mg daily.
4. Lantus insulin 50 units in the morning.
5. Lasix 80 mg daily.
6. Novolin R p.r.n.
7. Toprol XL 100 mg daily.
8. Flovent 100 mcg twice a day.

Diagnosis

1. Atherosclerotic coronary vascular disease with old myocardial infarction.
2. Moderate to severe LV systolic dysfunction.
3. Diabetes mellitus.
4. Diabetic nephropathy and renal failure.
5. Status post ICD implantation.

6. New onset of atrial fibrillation.
7. Chronic Coumadin therapy.

Plan

1. Continue present therapy.
2. Patient will be seen again in my office in four weeks

Box A

20 ways to misspell ““acetaminophen”

acetometaphin; acedeminifin; Acetemenophen; Acetimefin; Astemetaphine; Acetominephin; Metamaphine; Acetomenaphen; Acetominophin; Acetomenaphin; Acetometaphin; Achetophenomin; assidamidaphine; acetaminophens ; Acetamethane; Acetametaphin; Acetamedaphin; Acetominaphen; Acetimenophine, acetometaphin

HIGHLIGHTS

- taxonomy of errors of lay comprehension of medical documents is feasible
- this paper offers a classification serving as the beginning of such a taxonomy
- content analysis of document retellings is a viable taxonomy building procedure
- lay participants had the most trouble with medication names
- other difficulties concerned clinical concepts, terms, findings and research aims

Table 1

Characteristics of participants

Variable	Values
Gender (n, %)	
Female	64 (80)
Male	16 (20)
Age (years), n (%)	
<30	56 (70)
30–39	10 (12.5)
40–49	8(10)
50–65	4(5)
>65	1 (1.25)
No response	1 (1.25)
Education level attained[a], n (%)	
High school	10 (12.5)
College degree	46 (57.5)
Master's	21 (26.25)
>Master's	3 (3.75)
Degree type, n (%)	
Health-related	4 (5)
Nonhealth-related	71 (88.75)
No response	5 (6.25)
Biomedical knowledge[b]	
Mean (SD)	2.06 (1.06)
Diabetic knowledge[c]	
Mean (SD)	2.54 (1.25)

[a] of highest degree attained

[b] on scale from 1 ("I rarely read texts on biomedical topics") through 4 ("I read and understand general medical articles")

[c] on a scale from 1 ("very little") to 5 ("a good deal")

Table 2

Lay error classification scheme, on the basis of participants' retellings of two documents

Categories and Subcategories	Examples
Clinical concepts	
Clinical concept, incorrectly explained (CC) - *Incorrect explanation of a disease mechanism or biological process or concept*	"insulin is an enzyme"; "diabetes is a disease where the liver can't produce a certain type of sugar"
Clinical research	
Research purpose, inaccurate (RPI) *Misunderstanding the objective of the clinical trial, which is to develop a new method for assessing glycogen metabolism in patients with Type I Diabetes*	"the goal is to develop a new treatment for diabetes"; "this study develops a new method to diagnose diabetes
Medications	
Medication nature, incorrect (MedNI) - *Ascribing an incorrect function to a medication; calling a non medicinal substance a medicine*	"Coumadin is a sleep medication"; "2H2O is a medicine"
Medication, generic name misspelling - *Misspelling of a recognizable generic drug name*	"acet**o**minophin"
Medication, non-prescribed - *1) Non prescribed medication type, generic name, or form, 2) a medication purpose not corresponding to any medications in the documents or 3) a name that looks like a brand name, but isn't, and cannot be related to anything in the text*	"antibiotic," "aspirin," "some nasal spray," "medication for liver functioning," "Devton"
Medication, brand name misspelling - *A spelling error in a brand name that can be corrected by replacing or deleting one letter or switching two letters with each other*	*"ambion" or "ambient" for "ambient"; "Coudamin" for "Coumadin"*
Medication, brand name confusion - *"Swapping" two existing medications with each other; brand not mentioned in the document, but it does exist.*	*"Flomax" for "Flovent"*
Medication, partial memory - *More than 1 character is incorrect, but the original brand name mentioned in the document is generally recognizable or can be inferred*	*"Courdin" for "Coumadin," "Landin" for "Lantis"*
Medication units, incorrect - *Test results or medication dosage reported in the wrong units*	*respiration 18 **beats** per minute*
Medication regimen, incorrect - *Medication regimen / schedule/ dosage reported incorrectly*	*"as needed" instead of "daily" "1,000 mg" for "10 mg"*
Devices	
Device, incorrect explanation - *Using incorrect device name or misrepresenting its general purpose*	*"pacemaker" for "ICD implant," "spectroscopy machine for taking X-Rays"*
Procedures	
procedure, incorrect explanation - *Incorrect name or purpose , specific steps, or mechanism of a procedure*	*"detects low level of blood glucose" instead of "measures glycogen metabolism in the liver," "this method involves pairing a complex device with a painkiller," by means of "electromagnetic resonance" instead of "magnetic resonance"*
Terminology	
Clinical term, misspelling - *Misspelling of medical and health-related terms*	*"ateroscleriosis" for "atherosclerosis"*
Terminology confusion - *Substituting a specialized medical term with another medical or general term or a non-word that is similar to **another** medical or general term, confusing a medical term with a medication name, or making an error in an abbreviation or acronym; applying an inappropriate medical term to a contextually relevant description.*	"*diabetic phrenopathy" for "diabetic nephropathy," "Lasik" for "Lasix," ICV implant" for "ICD implant", "too much sugar in the blood is called "**hypoglycemia**'"*
Findings*	
Finding, inaccurately reported -*Misrepresenting some characteristic of a finding reported in the Visit Note, but clinically possible and the original source is clear*	*"hard abdomen" instead of "soft abdomen"*
Finding, non-existent - *Clinically possible, but no clear source in the Visit Note*	"frequent pain"

Categories and Subcategories	**Examples**
Clinical concepts	
Finding, nonsensical - *Clinically impossible finding*	*" trouble breathing a green expectorant"*
Diagnosis	
Diagnosis, wrong - *Attributing as a diagnosis a disease or a condition, not mentioned in the original document*	*"patient suffers from neuropathy"*
Diagnosis, inaccurately reported - *Misrepresenting some characteristics of a diagnosis mentioned in the document.*	*"acute diabetes" instead of "Diabetes Mellitus"*
Diagnosis, nonsensical - *Clinically impossible / meaningless diagnosis*	*"ischemic lung disease"*
Oher	
Non-existent direction - *Inaccurate recall of instructions regarding self-care and follow up visit*	*"return to the office* ***in a few months"***
Demographics, inaccurately reported - *Incorrect report of patients age / race*	*"60 - year old man"*
Patient's circumstances, inaccurately reported - *Inaccuracy in reporting the details of the patient's knowledge or time of treatment and visits*	*"patient was on a medication, but was not sure what it was treating"*

* Note: A finding is a clinically significant observation or measure (or set of observations or measures), potentially indicative of an underlying medical problem [67].

Table 3

Participant errors, by document and error category

Coding category	Subcategories	CT errors (# part)*	VN errors (# part)	Total Errors (# part)
Clin concepts	clinical concept, incorrectly explained (CC)	35 (25)	18 (15)	53 (33)
Clin research	research purpose, inaccurate (RPI)	42 (41)	0	42 (41)
Medications	Medication nature, incorrect (MedNI)	2 (2)	1 (1)	3 (3)
	Medication, generic name misspelling	19 (19)	1 (1)	20 (20)
	Medication, non-prescribed	3 (3)	14 (12)	17 (13)
	Medication, brand name misspelling	0	18 (17)	18 (17)
	Medication, brand name confusion	0	9 (8)	9 (8)
	Medication, partial memory	0	24 (16)	24 (16)
	Medication units, incorrect	0	1 (1)	1(1)
	Medication regimen, incorrect	0	1 (1)	1 (1)
	Total for Medications	**24(24)**	**69 (57)**	**93 (79)**
Devices	Device, incorrect explanation	1 (1)	5 (5)	6 (6)
Procedures	Procedure, incorrect explanation	12 (12)	0	12 (12)
Terminology	Terminology confusion	33 (27)	16 (15)	49 (36)
Findings	Finding, inaccurately reported	0	31 (25)	31 (25)
	Finding, non-existent	0	9 (8)	9 (8)
	Finding, non-sensical	0	12 (12)	12 (12)
	Total for Findings	**0**	**52 (45)**	**52 (45)**
Diagnosis	Diagnosis: wrong	0	4 (4)	4 (4)
	Diagnosis: inaccurately reported	0	7 (7)	7 (7)
	Diagnosis, non-sensical	0	1 (1)	1 (1)
	Total for Diagnosis	**0**	**12 (12)**	**12 (12)**
Oher	Non-existent direction	0	13 (13)	13 (13)
	Clinical term, misspelling	13 (8)	31 (20)	44 (25)
	Demographics, inaccurately reported	0	2 (2)	2 (2)
	Patient's circumstances, inaccurately reported	0	9 (8)	9 (8)

*The first number is the number of errors of a given type, the second number (in parentheses) indicates how many of the 80 participants made this type of error.