Balancing Structure and Narrative in Clinical Documentation

BMI 510: Introduction to Biomedical Informatics

Adam Rule

Introduction

One of the central challenges of clinical documentation is balancing structure and narrative.[1,2] Whereas narrative accounts concisely convey clinical information and reasoning, they are remarkably difficult to use for quality improvement or research due to their varying structure and representation of clinical concepts. Conversely while structured data such as diagnostic and procedure codes support computational analyses, they lack the context or nuance of narrative accounts which are vital to forming a coherent understanding of a patient's case. This tension between narrative and structure has become even more apparent with the recent shift in medical record use.[3,4] Whereas for millennia patient records were kept primary in the form of short educational narratives, over the past century they have gained structure to support billing, quality improvement, and research.

As medical records have become more structured, there has been growing discontent with their design and use. Clinical notes in particular are now commonly accused of being bloated and redundant, containing boilerplate text and tables to meet billing and regulatory requirements which obscure underlying narrative.[5–8] In an oft maligned example, the Evaluation and Management coding scheme for Medicare and Medicaid reimbursement stipulates how many systems need to be reviewed for different levels of payment, leading clinicians to include boilerplate text in their notes such as "ten-point review of systems returned negative" or lengthy tables of default exam findings.[9] Several organizations including the American Medical Informatics Association and the American College of Physicians have responded by releasing policy papers calling for change in how clinical documentation is regulated and performed with a renewed focus on capturing longitudinal histories to support patient care.[9,10]

This paper explores the challenge of balancing structure and narrative in clinical documentation. It starts with a brief history of the form and function medical records, discusses tradeoffs between structured and unstructured data, and concludes with an exploration of how health information technology (HIT) has been used and might be redesigned to balance structure and narrative in clinical documentation.

Brief History of Medical Records

Clinicians have kept medical records for thousands of years and the evolution of records over this time reflects the recurring challenge of balancing structure and narrative. The earliest known medical records are case books compiled in antiquity to track cases of educational value.[4] The Edwin Smith Papyrus written in roughly 1600 BC in Egypt, for example, contains 48 narrative cases discussing common injuries and how to treat them.[11] Other cases such as those contained in the Hippocratic Corpus focus on the stories of individual patients. One Hippocratic case, for example, opens with "the wife of Dromeades having been delivered of a female child, and all other matters going on properly, on the second day after was seized with rigor and acute fever...".[12] These accounts invite readers to reexperience the drama and uncertainty of the original case.

The use of case books for education continued relatively unchanged for centuries until the enlightenment when a renewed focus on natural science and principled observation encouraged an increasing number of physicians to keep personal collections of noteworthy cases.[4] Around the turn of the 19th century, the growing institutionalization of medicine and use of quantitative measurement to diagnose disease (rather than subjective symptoms) led to more disciplined capture of clinical data, often at the prompting of hospital administration.[4] Hospitals began to keep one-line records of every admission and discharge, then to select

noteworthy cases to transcribe from physicians' personal records into hospital case books, then to transcribe *all* cases into comprehensive case books with standardized sections for family history, previous illnesses, examination results, progress notes, and discharge instructions.[3] These reforms were not unopposed. For example, an 1885 proposal at the New York Hospital which suggested that "the original record of cases written at the bedside be made with such care and in such a manner that they may constitute the permanent records of the Hospital work" was not originally adopted due to the constraint providers felt it would put on their practice.[3] The policy was eventually adopted 15 years later after the burden of copying cases from provider's notes into hospital case books became too great. In the following years hospitals continued to impose additional structure on medical records, designing separate sheets to track orders, labs, and notes.[4] How records were cataloged began to change as well. For example, in 1914 the New York Hospital began to organize cases by disease and anatomical system to support identification of patient cohorts for research.[3]

Since case books were often organized by discharge date, ward, or disease, an individual patient's record might be scattered across multiple entries in multiple books, making it difficult to piece together a coherent story of their care. It was not until 1907 in the United States that patient records began to be organized around individual patients and kept at the bedside, due in large part to the efforts of Henry Plummer at the Mayo Clinic.[4,13] The New York Hospital and others adopted similar systems of organizing records by patient in the 1930s.[3]

Despite these efforts to reorganize records around the patient, it continued to be difficult to construct a clear narrative of an individual patient's care as data proliferated and spread across separate specialized sheets. This difficulty of reconstructing the longitudinal record, along with the advent of computerized patient records, led Larry Weed to start advocating for problem-

oriented medical records.[14] In a now famous grand rounds at Emory University in 1971 Dr. Weed pulled a record from one of Emory's clinics to demonstrate the difficulty of ascertaining the patient's problems, how they were treated, and why[15]. As he noted in the lecture, "you can't look at the management of a single problem without looking at the context, what are all the problems… And you shouldn't have to spend a second finding what are all the problems." While some of Dr. Weed's recommendations for the structure of medical records failed to catch on, his suggestion to organize notes into subjective, objective, assessment, and plan sections (i.e., SOAP notes) has now become standard practice in hospitals and medical schools around the world.[16]

The widespread adoption of electronic health records (EHRs) in the past decade has led to another shift in the way medical records are organized. In particular, researchers and clinicians note that EHRs have continued to fragment clinical data, organizing it by type of information (e.g., labs, medications, imaging) rather than by patient problem.[17] As a result, clinicians may still have difficulty piecing together a patients' story, even as they have ready access to more patient information than ever before. Moreover, increasing regulation about the design of EHRs and the content of notes has bloated clinical documentation with boilerplate text and tables that can obscure an underlying narrative.[6] This has led many to conclude that clinical notes have largely lost their narrative function.[9,10]

While reflecting the tension between narrative and structure, this history also shows how both policy and technology influence record keeping. In some cases, such as that of SOAP notes, a change of policy and social practice affects documentation without changing the underlying technology used to document (e.g., paper or computer). In other cases, the embedding of structure into technology (e.g., data entry forms, flow sheets) affects and enforces the change.

Tradeoffs of Structured and Unstructured Clinical Data

The tension between structured and unstructured data in medical records persists in large part because structured and unstructured data have different strengths.[1,2] Data entry, for example, can be completed more quickly when done in unstructured formats. In particular, studies have routinely shown that physicians take longer to draft electronic notes than paper ones.[18,19]. While it is unclear why this is the case – it could be due to initial unfamiliarity with electronic note entry or poor usability of note-entry interfaces – one likely factor is an expectation that electronic notes include more structured information such as review of systems or medication lists which are difficult to find and transcribe into notes manually. Since speed of note entry is a top priority for providers,[20] providers have adopted numerous methods to speed up electronic documentation including use of templates and text-importing phrases to automatically import structured information from other parts of the record.[6] Other studies more cleanly support the argument that structured data entry is less efficient by showing that clinicians take longer to place orders using computerized physician order entry than paper prescription pads.[21]

While potentially taking longer, structured data entry can lead to more complete data capture by requiring certain data such as primary diagnosis be captured for every patient visit.[22] However, other aspects of data quality such as the correctness of data are harder to assess.[23] There is limited evidence that data captured in free-text notes may actually be more accurate than data captured in face-sheets,[24] and growing concern that reliance on copy-paste to construct semi-structured notes is introducing errors that affect diagnosis and care.[25,26]

When it comes to using data, narrative can better support decision making and communication. This is in part due to the expressivity of narrative which can convey nuance such as the severity of disease, the provider's level of concern, and level of diagnostic

uncertainty more easily than structured terminologies.[1] Narrative also facilitates sensemaking by stringing together disparate pieces of information to make an argument. However, narrative is famously difficult to analyze, first requiring natural language processing to extract and harmonize data from across unstructured notes.[7,27] Structured data by contrast may require less cleaning and be easier to use for real-time decision support.

In the end, different uses of clinical data benefit from it being captured and stored in different formats. Unfortunately, data are not easily transformed from narrative into structed data or vice versa, requiring EHR designers to wrestle with how best to capture and display data to support the many uses of electronic records.

Balancing Narrative and Structure with Health Information Technology

One promising way to address the tension between structure and narrative in clinical documentation is to use health information technology to mix the entry and use of structured and unstructured data. One approach is to use natural language processing (NLP) to extract structured data from unstructured notes, either in real-time or after the fact. For example, the Medical Gopher system developed at the Regenstrief Institute periodically runs NLP on unstructured notes to do things like recognize medication orders and put links to these orders in a side-panel next to the note (Figure 1).[28] Outside clinical practice, numerous research projects make use to NLP to extract clinical data from unstructured notes.[27] A related emerging approach uses sensing in clinical environments to automatically detect procedures and then turn these into billing codes.[29]

A second approach to mixing structured and unstructured data requires clinicians to first input information in structured formats and then automatically generate free-text descriptions of that data to include in unstructured notes. For example, many EHRs include graphical user

interfaces which enable users to select common family history items or exam findings from a set of radio-buttons, checkboxes and dropdowns. This structured data entry can then be used to write formulaic text or generate tables which can be included in draft clinical notes. Other tools like note templates and smart phrases can be used to import boilerplate text and data from elsewhere in the record and similarly take structured data and turn it into unstructured note text.[6] Recent studies demonstrate that uses of these technologies is widespread and up to three-quarters of note text in some specialties may be constructed using these aids.[30]

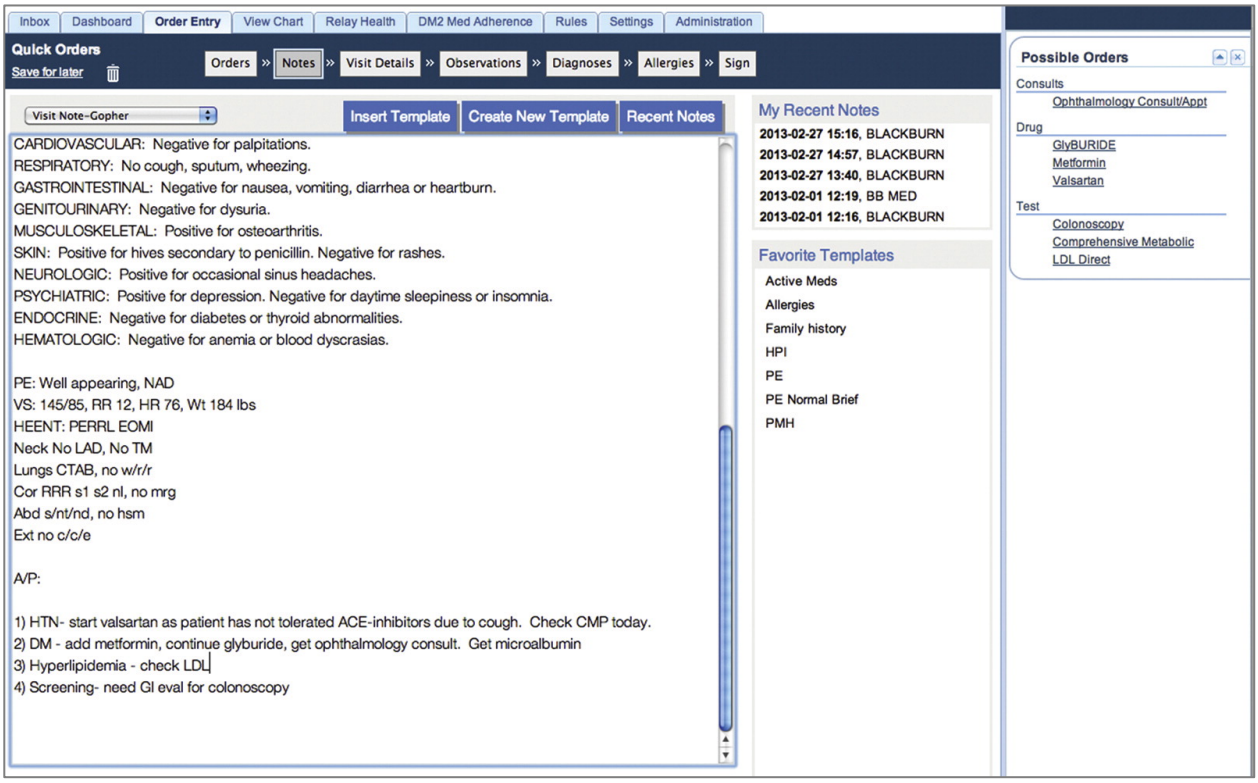

Figure 1: A recent update of the Medical Gopher EHR uses NLP to identify possible orders in the assessment and plan sections of draft notes and populates a side-panel with links to place these orders.[28]

One shortcoming of both these approaches is that they may break the link between structured and unstructured data once converted. Once free-text has been parsed to generate an order, it may not be possible to work backwards from that order to determine which snippet of

free text was used to generate it. Similarly, once structured data has been turned into free-text via a note writer, template, or smart phrase, it may not be possible to work back from the free text to view the underlying data point. This introduces redundancy in the record and potential for unstructured and structured representations of the same data-point to conflict.

A third approach to mixing structured data and narrative seeks to address this issue by linking structured and unstructured data more closely in the same document, enabling clinicians to simultaneously work with both representations. For example, Johnson et al. at Columbia prototyped an EHR based on structured narrative which used XML documents to markup free text with metadata such as document sections, clinical concepts, and relationships between them.[2] The authors speculate how this approach might be used to ease the process of reviewing batches of clinical notes, or to extract data to enable decision support and research. Wilcox et al, also at Columbia, later experimented with a system that automatically parses free text within a note to search for related information in other parts of the record (Figure 2).[31] In the case of a provider searching for the concept “ABP”, the system would show related information in a side-panel such as a graph of arterial blood pressure over time or notes mentioning the concept. It would then enable users to quickly copy that information into their draft note. Critically, this information was not just copied but rather transcluded with a link back to the original source of the information. Clinicians could even ask to be notified when the transcluded information changed, turning the note into a preemptive paging system. These systems build on the insight that much of note writing is actually data synthesis, looking for information in other parts of the record and combining it with fresh findings to create a coherent story.[32] Rather than do this consolidation automatically via a template, they engage clinicians

in the sensemaking process, potentially improving comprehension and limiting the amount of extraneous information included in notes.

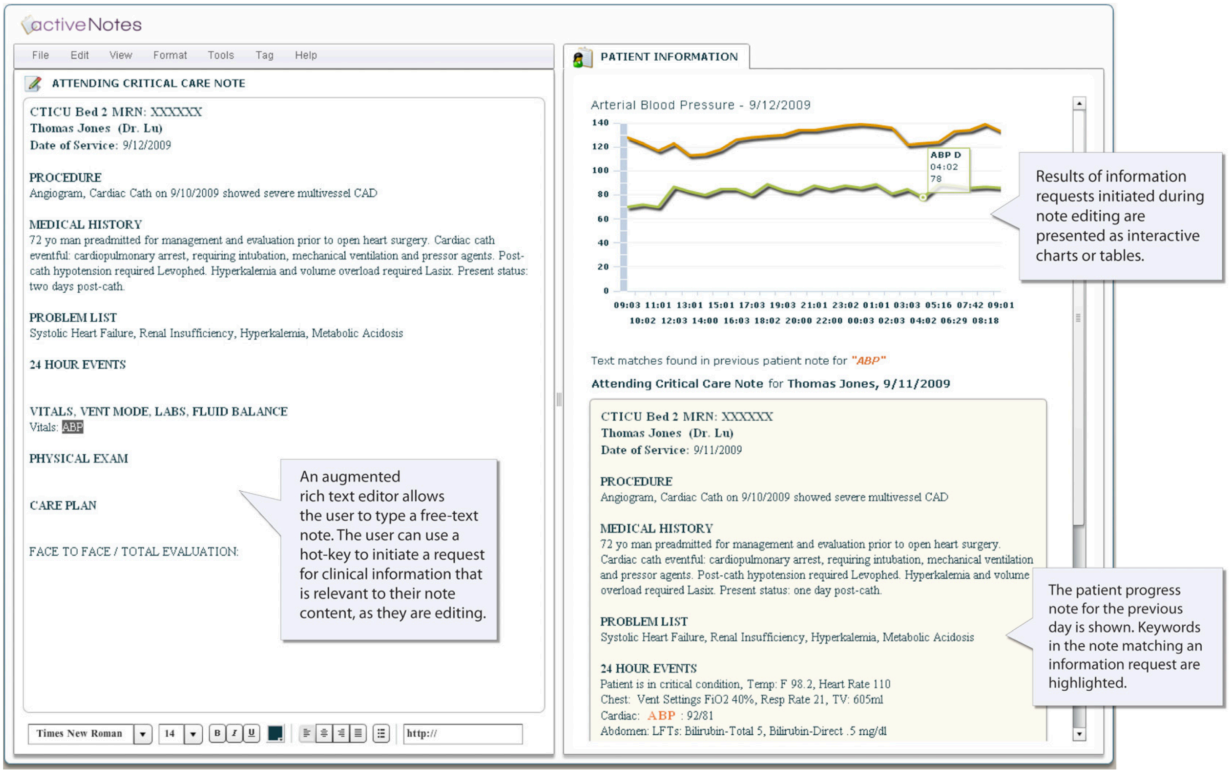

Figure 2: The activeNotes system enabled clinicians to search for data using note text, view that information in a side panel, and include relevant content back in their draft note.[31]

This final approach of mixing structured and unstructured data in the same document holds promise but adds complexity to how data are represented and may place additional burden on providers to ensure that free-text is parsed into structured data correctly. However, it may also reduce the need to perform an action (e.g., place an order) and then separately document that action in an unstructured note. Rule et al. previously explored this possibility in a prototype note-editor developed with the Veterans Medical Research Foundation that let clinicians place medication orders by writing free text into a search box that would support search and autocompletion of medication orders based on a formulary, place a textual representation of that

order inline in the note, and prepare a structured version of that order for the provider to sign.[33] Critically, the free-text representation of the order and the structured representation were closely linked; editing the text in the note would edit the order. While this was a prototype system, it demonstrates an approach to drafting medical records that seeks to unify action and documentation. This approach might be fleshed out into more comprehensive notebooks, such as those used for data analysis, which enable providers to search for and visualize information, perform actions such as placing orders, and annotate both with free text explaining clinical reasoning.

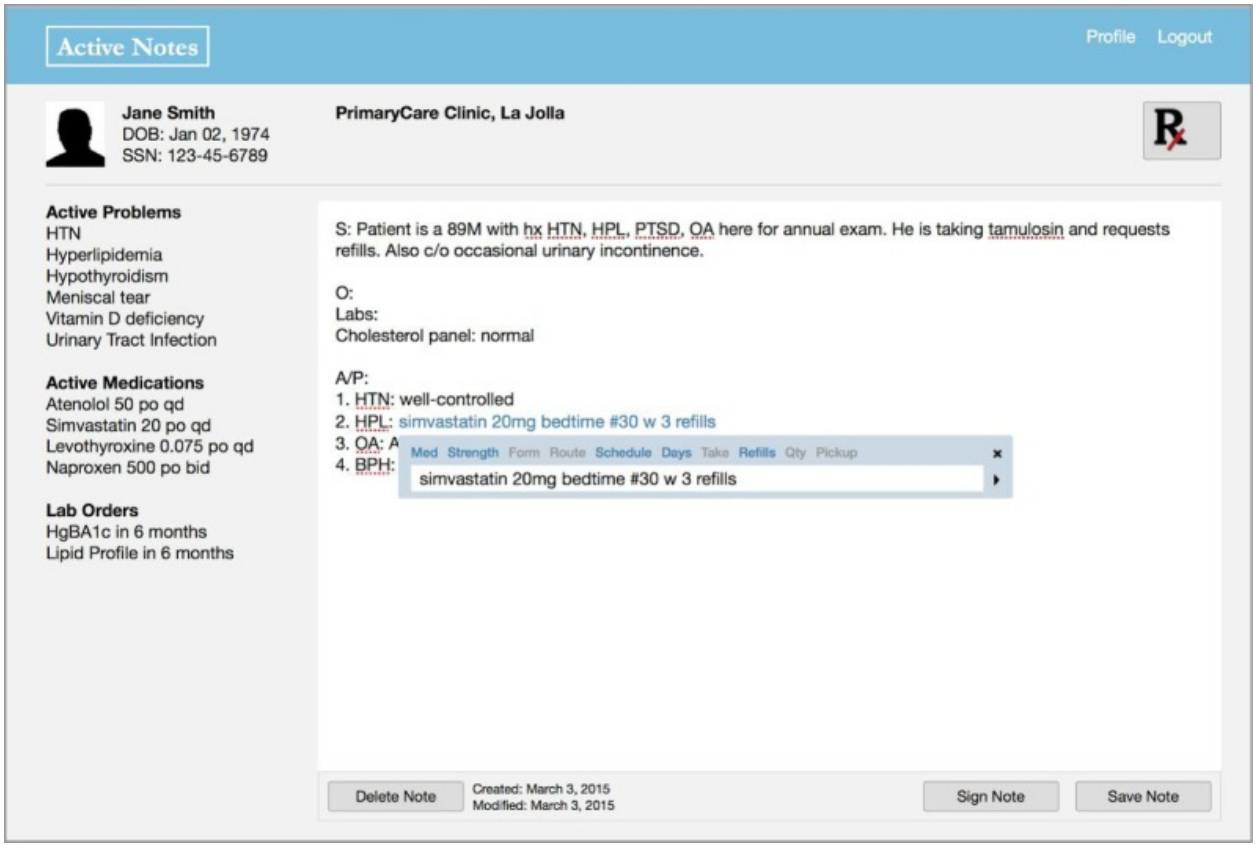

Figure 3: The ActiveNotes system enabled clinicians to place inline-medication orders using a search bar placed text into the note and linked this to a structured medication order for the provider to sign.[33]

Conclusion

The history and current design of medical records reflects a tension between narrative and structured data. Over time, medical records have transitioned from being largely unstructured narrative accounts of patient cases used didactically to increasingly structured representations of patient care that can be analyzed at scale. While narrative better supports communication and sensemaking, structured data is more readily analyzed to support monitoring and knowledge discovery. While there has historically been a divide between structured and unstructured representations of the same clinical event, advances in data structures and computer interfaces are enabling clinicians to work with clinical documents that increasingly mix structure and narrative. Development of rich notebooks mixing information retrieval, visualization, order entry, and documentation might relieve some of the tension between narrative and structure, enabling documentation to better serve clinicians' data and narrative-driven workflows.

References

1 Rosenbloom ST, Denny JC, Xu H, *et al.* Data from clinical notes: a perspective on the tension between structure and flexible documentation. *J Am Med Inform Assoc JAMIA* 2011;**18**:181–6. doi:10.1136/jamia.2010.007237

2 Johnson SB, Bakken S, Dine D, *et al.* An electronic health record based on structured narrative. *J Am Med Inform Assoc JAMIA* 2008;**15**:54–64. doi:10.1197/jamia.M2131

3 Engle RL. The evolution, uses, and present problems of the patient's medical record as exemplified by the records of the New York Hospital from 1793 to the present. *Trans Am Clin Climatol Assoc* 1991;**102**:182–9; discussion 189-192.

4 Gillum RF. From papyrus to the electronic tablet: a brief history of the clinical medical record with lessons for the digital age. *Am J Med* 2013;**126**:853–7. doi:10.1016/j.amjmed.2013.03.024

5 Hirsch AG, Jones JB, Lerch VR, *et al.* The electronic health record audit file: the patient is waiting. *J Am Med Inform Assoc JAMIA* 2017;**24**:e28–34. doi:10.1093/jamia/ocw088

6 Weis JM, Levy PC. Copy, Paste, and Cloned Notes in Electronic Health Records. *Chest* 2014;**145**:632–8. doi:10.1378/chest.13-0886

7 Cohen R, Elhadad M, Elhadad N. Redundancy in electronic health record corpora: analysis, impact on text mining performance and mitigation strategies. *BMC Bioinformatics* 2013;**14**. doi:10.1186/1471-2105-14-10

8 Wrenn JO, Stein DM, Bakken S, *et al.* Quantifying clinical narrative redundancy in an electronic health record. *J Am Med Inform Assoc* 2010;**17**:49–53. doi:10.1197/jamia.M3390

9 Kuhn T, Basch P, Barr M, *et al.* Clinical documentation in the 21st century: executive summary of a policy position paper from the American College of Physicians. *Ann Intern Med* 2015;**162**:301–3. doi:10.7326/M14-2128

10 Cusack CM, Hripcsak G, Bloomrosen M, *et al.* The future state of clinical data capture and documentation: a report from AMIA's 2011 Policy Meeting. *J Am Med Inform Assoc JAMIA* 2013;**20**:134–40. doi:10.1136/amiajnl-2012-001093

11 Nissen T, Wynn R. The history of the case report: a selective review. *JRSM Open* 2014;**5**:205427041452341. doi:10.1177/2054270414523410

12 Hippocrates. *On the Epidemics*. http://classics.mit.edu/Hippocrates/epidemics.html (accessed 16 Aug 2019).

13 Camp CL, Smoot RL, Kolettis TN, *et al.* Patient records at Mayo Clinic: lessons learned from the first 100 patients in Dr Henry S. Plummer's dossier model. *Mayo Clin Proc* 2008;**83**:1396–9. doi:10.1016/S0025-6196(11)60790-8

14 Weed LL. Medical Records That Guide and Teach. *N Engl J Med* 1968;**278**:652–7. doi:10.1056/NEJM196803212781204

15 Larry Weed's 1971 Internal Medicine Grand Rounds. 1971. https://www.youtube.com/watch?v=qMsPXSMTpFI (accessed 23 Aug 2019).

16 Wright A, Sittig DF, McGowan J, *et al.* Bringing science to medicine: an interview with Larry Weed, inventor of the problem-oriented medical record. *J Am Med Inform Assoc JAMIA* 2014;**21**:964–8. doi:10.1136/amiajnl-2014-002776

17 Heath C, Luff P. Documents and professional practice: "bad" organisational reasons for "good" clinical records. In: *Proceedings of the 1996 ACM conference on Computer supported cooperative work - CSCW '96*. Boston, Massachusetts, United States: : ACM Press 1996. 354–63. doi:10.1145/240080.240342

18 Chaudhry B, Wang J, Wu S, *et al.* Systematic review: impact of health information technology on quality, efficiency, and costs of medical care. *Ann Intern Med* 2006;**144**:742–52. doi:10.7326/0003-4819-144-10-200605160-00125

19 Jones SS, Rudin RS, Perry T, *et al.* Health information technology: an updated systematic review with a focus on meaningful use. *Ann Intern Med* 2014;**160**:48–54. doi:10.7326/M13-1531

20 Payne TH, tenBroek AE, Fletcher GS, *et al.* Transition from paper to electronic inpatient physician notes. *J Am Med Inform Assoc* 2010;**17**:108–11. doi:10.1197/jamia.M3173

21 Bates DW, Boyle DL, Teich JM. Impact of computerized physician order entry on physician time. *Proc Symp Comput Appl Med Care* 1994;:996.

22 Bell DS, Greenes RA. Evaluation of UltraSTAR: performance of a collaborative structured data entry system. *Proc Symp Comput Appl Med Care* 1994;:216–22.

23 Weiskopf NG, Weng C. Methods and dimensions of electronic health record data quality assessment: enabling reuse for clinical research. *J Am Med Inform Assoc JAMIA* 2013;**20**:144–51. doi:10.1136/amiajnl-2011-000681

24 Fox KM, Reuland M, Hawkes WG, *et al.* Accuracy of medical records in hip fracture. *J Am Geriatr Soc* 1998;**46**:745–50. doi:10.1111/j.1532-5415.1998.tb03810.x

25 Weir CR, Hurdle JF, Felgar MA, *et al.* Direct text entry in electronic progress notes. An evaluation of input errors. *Methods Inf Med* 2003;**42**:61–7.

26 Singh H, Giardina TD, Meyer AND, *et al.* Types and origins of diagnostic errors in primary care settings. *JAMA Intern Med* 2013;**173**:418–25. doi:10.1001/jamainternmed.2013.2777

27 Kreimeyer K, Foster M, Pandey A, *et al.* Natural language processing systems for capturing and standardizing unstructured clinical information: A systematic review. *J Biomed Inform* 2017;**73**:14–29. doi:10.1016/j.jbi.2017.07.012

28 Duke JD, Morea J, Mamlin B, *et al.* Regenstrief Institute's Medical Gopher: a next-generation homegrown electronic medical record system. *Int J Med Inf* 2014;**83**:170–9. doi:10.1016/j.ijmedinf.2013.11.004

29 Farooq F, Overhage J. Systems and methods for capturing data, creating billable information and outputting billable information. 2016.

30 Hribar MR, Rule A, Huang AE, *et al.* Redundancy of Progress Notes for Serial Office Visits. *Ophthalmology* Published Online First: 21 June 2019. doi:10.1016/j.ophtha.2019.06.015

31 Wilcox L, Lu J, Lai J, *et al.* Physician-Driven Management of Patient Progress Notes in an Intensive Care Unit. *Proc SIGCHI Conf Hum Factors Comput Syst CHI Conf* 2010;**2010**:1879–88. doi:10.1145/1753326.1753609

32 Mamykina L, Vawdrey DK, Stetson PD, *et al.* Clinical documentation: composition or synthesis? *J Am Med Inform Assoc* 2012;**19**:1025–31. doi:10.1136/amiajnl-2012-000901

33 Rule A, Rick S, Chiu M, *et al.* Validating free-text order entry for a note-centric EHR. *AMIA Annu Symp Proc AMIA Symp* 2015;**2015**:1103–10.