

Human-AI Collaboration “In the Wild”: The Case of Clinical Documentation Integrity Specialists

Kathleen H. Pine

Arizona State University, khpine@asu.edu

Claus Bossen

Aarhus University, clausbossen@cc.au.dk

This paper describes the work of clinical documentation integrity specialists (CDIS), who review and code patient records to improve the quality of documentation by clinicians in the medical record concurrently with a patient's stay in the hospital. CDIS are a relatively new occupation whose number has grown in the 2010s with the introduction of the Affordable Care Act in 2010, as well as with the introduction of the more complex ICD-10 classification of diseases. Both require more accurate and fine-grained coding and have led to the implementation of 'computer-aided-coding' system (CAC). CDIS work with two computer systems: the electronic health record (EHR) and the CAC, which that applies natural language processing and statistics to scan records and suggest auto-generated codes for CDIS to evaluate and apply with discretion. CDIS are hence one example of human-AI collaboration in healthcare, which we can use to think about implications for the study of human-centered AI in HCI.

CCS CONCEPTS • Human-centered computing~Human computer interaction (HCI)

Additional Keywords and Phrases: Clinical documentation improvement specialists; computer assisted coding system; healthcare; AI

1 INTRODUCTION

Our in-progress research investigates the work of clinical documentation integrity specialists (CDIS) and how they work with AI “in the wild” as they work to improve healthcare coding. CDIS are an occupation that carries out the concrete work of producing data that feeds into the larger and more abstract process of datafication of healthcare. Their job is to audit clinician charting in close to “real time”—while patients are still in the hospital or soon after discharge—and query clinicians when they detect documentation that is inconsistent, lacks precision or missing information. One important concern in addition to documentation integrity is whether such deficiencies produce a less than optimal outcome ICD-10 code or DRG group given the patient's condition. Thus, CDIS sit “between” the long-established roles of clinicians and medical records ‘coders’. CDIS must have a clinical degree and the authority and legitimacy to assess medical documentation and send “queries” to clinicians asking for clarifications in the clinical documentation that may alter the code set applied to the medical record. Historically, this process primarily aimed at improving reimbursement from public and private insurance

for high-value charts through thoroughly capturing the severity of diseases and appropriateness of treatment and procedures. However, with the growth of national quality reporting and “report cards” and the move towards value-based reimbursement (also called “pay-for-performance”) programs for hospitals and clinicians (Pine & Mazmanian, 2015), CDIS are seen by many healthcare organizations as a necessary component of a strategy to produce data that will stand up to the rigors of measurement-based accountability initiatives (Lopez, 2010; Payne, 2015).

In our field sites, coders enter the final code set based on the coding of CDIS and their own reading of records, and submit these final codes to the billing department who use the code sets to bill public health insurance programs, private insurance companies, or the patients themselves. While coders have the ability to query clinicians directly, queries from coders to clinicians are exceedingly rare, since clinicians tend to feel that since coders have no clinical degree they have no place querying them. Also, since medical coders’ coding takes place after medical care is complete (and often has been complete for some time) clinicians often do not remember the details of a case precisely anyway. CDIS, on the other hand, have a clinical background and query clinicians often and are expected to do so. Their role, as the job title says, is to improve clinical documentation, before the coders get in the loop.

Prior to the introduction of electronic health records (EHR) at the turn of the century, the occupational role of CDIS was difficult if not impossible, because the paper-based medical record was too in-demand by clinicians to transfer it to an employee, who could carry out an entire coding process and monitor the evolution of clinical documentation in near real time. Implementation of EHRs afforded this new role as EHRs made it possible for multiple employees to access, read, and work with a patient’s medical record at the same time, without crowding medical units or interfering with clinical work.

Digitization of health records made auditing clinical documentation via coding concurrently with delivering clinical care possible, but the Affordable Care Act of 2010 provided a major push towards development of the occupation: The act required more detailed coding and laid out a series of steps leading toward value-based reimbursement, which meant that certain providers would be paid less, if they did not show quantitative evidence of providing high-quality care. These calculations are based on standardized codes applied to medical records, increasing the importance of these codes. Further, the implementation of ICD-10 (International Classification of Diseases) in 2015 expanded the ICD from 13,000 to 68,000 codes and added significantly to the complexity and level of specification in the ICD (Cartwright, 2013). Healthcare providers were required to code more and more accurately, which led to the rapid creation and expansion of the CDIS occupation as well as to the introduction of computer-assisted-coding-systems (CACS) that use statistics and natural language processing to auto-generate codes for CDIS.

2 THE WORK OF CDIS

As stated, the work of CDIS is to make sure that clinical documentation by doctors is precise, correct and comprehensive. Doctors are – for good reasons, CDIS acknowledge – primarily focused on keeping up with treating incoming patients. Spending time on documentation in order for records to be complete and accurate is a secondary concern. Especially, if this is for post hoc purposes such as billing and calculating quality measurements. However, from the perspective of healthcare service providers, incomplete records mean that hospitals may not be paid for the services they provide, resulting in economic deficits and hence potentially lower quality of treatment and care. On the other hand, CDIS must also avoid “over coding” that could lead to

over-billing and in turn trigger an audit of a provider. Hence, the task of CDIS is to ensure that that patient records are accurate and comprehensive, and coded according to given regulations.

Working as a CDIS hence includes making queries to doctors about the record and often asking them to be more specific. Did the patient have 'mild sepsis' or 'severe sepsis' at admittance? What kind of pneumonia did the patient have? These queries called 'Clinical clarifications' are conducted via special forms in the CACS. Coding in general requires medical knowledge by CDIS and most CDIS have worked as registered nurses for 10 years or more before they become CDIS. Especially with 'Clinical Clarification' is it important for CDIS to have clinical expertise and credentials, since their expertise helps to ameliorate animosity of clinicians at having non-clinicians question their charting and make the CDIS harder to ignore. Per regulation, clinical clarifications cannot be leading: CDIS must merely clarify language, not make suggestions that pressure or lead doctors to choose diagnoses or terms that result in higher reimbursement.

3 WORKING WITH THE CAC

Our CDIS, drawn from two different healthcare system field sites, work away from patients distributed across the healthcare network. In both healthcare system, CDIS divide their time between a basement offices with individual cubicles at the network and days spent working from home. All CDIS in both field sites worked with two monitor screens on which they could display data from two computer systems: One is the hospital's EHR in which they look at a particular patient's medical notes, examinations, test results, etc., and the other screen displays the CAC. The CAC processes the text in the EHR and autosuggests a number of codes that might apply to each patient record.

The CDIS are required to open and run the CAC while they code, but beyond this requirement the CDIS are free to work with the CAC as they like. Some preferred to code a patient's record first by going through physician and nurse notes, examinations, etc., and then consult the CAC afterwards. Others start out from the CAC's suggested codes and go through these to see whether they are relevant or not. For each code suggested by the CAC, the system will point to the particular part of the record, which has triggered the suggestion of a code. In this observation, a CDIS has already done a review of the patient record and made extensive notes, and is using the CAC to code by accepting, rejecting, or elaborating on the suggested codes (or entering new codes not suggested by the CAC):

"CDIS: Now, I've gone through, and I've taken everything here.

Interviewer: From the CAC.

CDIS: From the CAC. Now, I'm gonna go through each one of 'em. If it's just an R, that's just basically—it's found in the record. I can validate if it's true or not [using an inspection tool]. [The CAC has suggested a code for dementia that the CDIS is considering]. I have no information that she has dementia. Now, I'm gonna [click the inspection tool] --show me where you read that...I'm not gonna fly with that one yet. When I looked at the nutrition record, it doesn't sound like it's poor—it is poor PO intake ["per os", by mouth], but I can't say, because she's poor PO intake, that that's dementia. I don't see anybody else sayin' that. I'm gonna leave that one alone. I just take 'em off my list [by rejecting the code suggested by the CAC], so I know. [Seeing a code for benzo-dependence that she had not written in her own notes but knows is correct when she sees it] I guess I forgot to check this one. She's a benzo-dependent person, so I'm gonna select that one."

As we can see in this example, CDIS are not required to make use of the auto-suggested code, or account for their use or non-use of the suggested codes—they can simply reject almost all suggested codes. Rather, the criteria of successful coding is the degree to which their coding is challenged by the coder or the insurance company, or whether their Clinical Clarification led to a code resulting in higher reimbursement. The freedom to work with the auto-suggested codes in different ways – indeed even to completely ignore them – makes clear the shortcomings of the technology. The CDIS we interviewed and observed are themselves quite skeptical towards the CAC: They feel it suggests many obviously irrelevant codes, and often the codes indicate a ‘stupid’ machine reading of text. Another popular example was when the CAC suggested a code for ‘trauma’, because a record included the statement by a patient that ‘I felt like I had been kicked by a horse’ in response to new information from the physician. This example and others contributed to CDIS treating auto-suggested codes with suspicion and only accepting them once a careful check in the record had been conducted.

Interestingly, even though the CDIS treat the codes generated by the CAC with caution and even disdain, the fact that they are required to work with the CAC implies that CDIS managers attribute some amount of authority to the AI system. Similarly to the concept of “algorithmic authority” whereby algorithms are granted some authority simply by account of the fact that they are algorithms and perceived as “objective” and “impartial” thus superior actants to humans (Lustig & Nardi, 2015), the AI of the CAC system is granted authority simply by virtue of being not-human, thus CDIS are required to deploy it even when they have direct evidence that the AI makes even obvious errors assembling accurate code sets. Although we expected some CDIS to be frustrated by the fact that they cannot turn the CAC off thus are required to engage with it, CDIS told us that they find the CAC helpful despite its flaws. As one CDIS evocatively described:

“[The CAC] helps me because sometimes the human eye—we have error, and I think together we’re a good team. That’s what I will say. It’s like we’re Batman, and [the CAC] is Robin, and it’s definitely not the Batman because it doesn’t have a real brain to know. I mean, it makes a lotta mistakes, but we all—we’re human. I could be sick. I could have allergies, and I could be rubbin’ my eyes and tryin’ to do my work, and I’m like, “I missed that, and [the CAC] got it,” and then [the CAC], sometimes it’ll not pick up stuff that I caught, and I’m like, “I got it.” I think it’s good teamwork. It’s just having an annoying partner that doesn’t always do what it’s supposed to, or it doesn’t—what it’s intended to do, it might be slow-functioning that day, or it’s having glitches that day. It’s like sometimes having a partner that you’re workin’ with that’s sick.”

As this CDIS describes, the CAC is helpful even though it is often not very good. CDIS describe the CAC in a variety of such terms—a sort of inept helper, but one that nonetheless adds enough value that CDIS prefer to have it around.

Another element at play is that CAC systems are learning systems that refine their text mining with ongoing use, so in a sense the CDIS who use them are “teaching” the clunky systems as they adopt or discard certain codes suggested by the CAC. This raises an interesting human/AI interaction whereby the human must work with an AI to teach it, an act pervaded by the vague and abstract hope that perhaps the CAC will improve and over time the CAC will perform better for CDIS as a group. However, this also raises questions about the future role of the CDIS as the occupation evolves in relation to their AI tool. Another facet of working with the AI related

to learning and teaching is that the CDIS in our field sites are careful to not introduce the CAC into training of new CDIS until they are very comfortable with reviewing charts and building code sets without the CAC. CDIS deliberately limit exposure to AI in early training to ensure they do not become overly reliant on the CAC, and also to help build skill in discernment necessary to determine which auto-generated codes are good and which should be ignored. Working with the CAC is skillful work, and we plan to investigate this work further as a case of algorithmic competence (Jarrahi & Sutherland, 2019), that is, the expertise required to work with, against, or around AI as part of work practice.

4 WORKING WITH AI

In our next round of research, we want to look closer into the CDIS-CACS collaboration. Presently, the cooperation leaves autonomy and decision-making to the CDIS even though the CACS in question is marketed as the most advanced and state-of-art system. This autonomy granted to the CDIS is surprising since natural language processing based on text is a well-known challenge for AI, and also because these texts derive from medicine, which is a field with highly standardized vocabulary.

We propose that CDIS are an interesting case to investigate human-AI collaboration, and would like to pursue the following questions:

- Which specific challenges arise for CDIS when working with CACs?
- How do CDIS maneuver between the supposed abilities of CAC systems and their own expertise?
- How do CDIS conceptualize their role with relation to the CACS —e.g. what are their thoughts and feelings about “training” the CAC systems

For the workshop, we would like to contribute with a pertinent case of healthcare AI in the wild. From the workshop, we hope learn and contribute to conversations about how to go the “last mile” and develop human-centered AI systems that support healthcare practice and workers.

REFERENCES

- [1] Cartwright, D. J. (2013). ICD-9-CM to ICD-10-CM Codes: What? Why? How?. *Advances in Wound Care*, 2(10), 588.
- [2] Jarrahi, M. H., & Sutherland, W. (2019). Algorithmic management and algorithmic competencies: Understanding and appropriating algorithms in gig work. In *International Conference on Information* (pp. 578-589). Springer, Cham.
- [3] Lopez, B. A. (2010). Establishing a CDI program: How one organization leveraged nursing and coding skills to improve clinical data. *Journal of AHIMA*, 81(7), 58-59.
- [4] Lustig, C., & Nardi, B. (2015, January). Algorithmic authority: The case of Bitcoin. In *2015 48th Hawaii International Conference on System Sciences* (pp. 743-752). IEEE.
- [5] Payne, T. (2010). Improving clinical documentation in an EMR world: one health system used a three-step process to integrate a clinical documentation improvement program with its new electronic medical record system, with significant results. *Healthcare Financial Management*, 64(2), 70-75.
- [6] Pine, K., & Mazmanian, M. (2015). Emerging insights on building infrastructure for data-driven transparency and accountability of organizations. *iConference 2015 Proceedings*.