AMIA

INFORMATICS PROFESSIONALS. LEADING THE WAY.

OXFORD

Research and Applications

Characterizing outpatient problem list completeness and duplications in the electronic health record

Edward Chia-Heng Wang[1] and Adam Wright[2,3]

[1]Department of Medicine, Harvard Medical School, Harvard University, Boston, Massachusetts, USA, [2]Division of General Internal Medicine, Brigham and Women's Hospital, Boston, Massachusetts, USA and [3]Department of Biomedical Informatics, Vanderbilt University Medical Center, Nashville, Tennessee, USA

Corresponding Author: Adam Wright, PhD, Department of Biomedical Informatics, Vanderbilt University Medical Center, 2525 West End Avenue, Suite 1475, Room 14109, Nashville, TN 37203, USA; adam.wright@vumc.org

ABSTRACT

Objective: The study sought to characterize rates of problem list completeness and duplications in common chronic diseases and to identify any relationships that they may have with respect to disease type, demographics, and disease severity.

Materials and Methods: We performed a retrospective analysis of electronic health record data from Partners HealthCare. We selected 8 common chronic diseases and identified patients with each of those diseases. We then analyzed each patient's problem list for completeness and duplications and also collected information regarding demographics and disease severity. Rates of completeness and duplications were calculated for each disease and compared according to disease type, demographics, and disease severity.

Results: A total of 327 695 unique patients and 383 404 problem list entries were identified. Problem list completeness varied from 72.9% in hypertension to 93.5% in asthma, whereas problem list duplications varied from 4.8% in hypertension to 28.2% in diabetes. There was a variable relationship between demographic factors and rates of completeness and duplication. Rates of completeness were positively correlated with disease severity for most diseases. Rates of duplication were consistently positively correlated with disease severity.

Conclusions: Incompleteness and duplications are both important issues in problem lists. These issues vary widely across different diseases and can also be impacted by patient demographics and disease severity. Further studies are needed to investigate the effect of individual user behaviors and organizational policies on problem list utilization, which will aid the development of interventions that improve the utility of problem lists.

Key words: electronic health records, electronic medical records, problem lists

INTRODUCTION

Electronic health records (EHRs) play a central role in modern healthcare systems. They perform a multitude of functions, such as storing patient health records, recording clinical encounters, facilitating communication, managing orders, and serving as databases for research. Although EHRs have improved many workflows in medicine, such as the safety of medication orders and the practice of evidence-based medicine, there remains an immense opportunity for improvement.[1,2] In a 2015 consensus statement, the American Medical Informatics Association (AMIA) EHR-2020 Task Force identified increased data entry burden, misguided regulation, poor usability, and lack of interoperability as some of the most pressing issues facing modern EHRs.[3] These shortcomings have been linked with higher rates of burnout, decreased face-to-face time with patients, and lower job satisfaction among physicians.[4–7]

These shortcomings can have clinical implications as well. In particular, the clinical documentation of a patient's medical history is often scattered among hundreds, if not thousands, of separate clinical notes, making it difficult to understand the underlying medical narrative that is instrumental to medical decision making.[8] Problem lists proactively address this issue by summarizing each patient's medical problems in a centralized location, as opposed to reactive solutions such as search functions that simply help providers to navigate scattered information. Although ineffective problem lists are not among the most pressing issues identified by the AMIA EHR-2020 Task Force, understanding problem list completeness and duplications is essential to reduce the negative impact of incomplete and redundant problem lists.

Problem lists are essential to clinician workflows because they highlight critical problems that otherwise might be forgotten and make it easier for multiple providers to coordinate the management of each of those problems.[9] Similarly, problem lists make it easier for clinicians who are taking care of new patients to familiarize themselves with a patient's critical problems quickly and efficiently. This is especially important in fast-paced environments like the emergency department, where an in-depth review of the medical record is not always practical.[10]

Maintaining complete problem lists is particularly important because it can improve patient care in tangible ways, especially with tasks that are more susceptible to lapses in memory; these include the consistent prescribing of cardiovascular medications in patients with heart failure and the thorough follow-up of medical issues in the ambulatory setting.[9,11] Incomplete problem lists also engender distrust in the overall reliability of problem lists. This can lead to a "tragedy of the commons" situation in which providers react to incomplete problem lists by refusing to use or update them, which only exacerbates the problem.[12] Despite their importance, in practice, problem lists are often incomplete, redundant, or even inaccurate.[13,14] Furthermore, they often fail to document critical semantic information, beyond diagnostic codes, that is essential to medical decision making.[15] Without automated, easy-to-use tools to maintain problem lists efficiently, providers are left to manage them manually, which is often unfeasible. Consequently, the clinical utility of problem lists often languishes over time as they remain incomplete and accumulate duplicates.

Previous studies of problem list utilization have typically focused on problem list completeness. One study found that problem list completeness for patients with diabetes ranged from as little as 60.2% to 99.4%, depending on the presence of "positive deviance" factors, such as financial incentives.[16] Such stark differences in problem list completeness may also be related to healthcare providers' diverse attitudes toward issues such as problem list ownership and content.[12,17,18] From an interventional perspective, clinical decision support tools, automated alerts, and machine learning may play a role in improving problem list completeness across a wide range of diseases.[19–21]

Problem list duplications are an equally important issue. They are essentially an analogous manifestation of "note bloat," which has been faulted by the AMIA EHR-2020 Task Force for filling clinical documentation with distractors and often redundant, unreadable, or unreliable information.[3,22–24] Duplications compromise the fundamental purpose of problem lists: to provide a comprehensive, yet palatable, medical overview of a patient. Although some duplications are useful in explicitly identifying disease attributes such as complications and stages, they often simply present redundant information multiple times. At best, these redundancies are purely visual distractions that slow providers down, but at their worst, they can become so visually and cognitively burdensome that they obscure other crucial health issues from providers. Redundant entries can also cause confusion through inconsistencies, and they can interfere with problem-based charting by providing multiple, disparate channels for the documentation of a single issue. Thus, resolving duplications is key to improving problem lists, but this issue has not been well studied in the literature.

Similarly, little is known about the effect of patient factors, such as demographics and disease severity, on problem list utilization. This information can help focus future EHR interventions on specific patient populations that are most vulnerable to the aforementioned negative effects of incomplete and duplicated problem lists. We hypothesize that problem list completeness improves for patients who fit classic presentations of disease and who have more severe disease. If this is truly the case, interventions should reduce this bias so that more patients with all kinds of presentations and disease states can have a complete problem list. We also hypothesize that problem list duplications increase for patients with high healthcare utilization and more severe disease, so future interventions should reduce duplications for these particular patients.

MATERIALS AND METHODS

Data source and patient identification
We conducted a retrospective analysis of EHR data from Partners HealthCare, which uses Epic Systems Corporation (Verona, WI) as its commercial EHR vendor. We focused on 8 common chronic diseases: Crohn's disease, ulcerative colitis, depression, schizophrenia, diabetes, hypertension, asthma, and epilepsy. For each of these diseases, diagnostic criteria were developed, as detailed in Supplementary Table 1. These criteria were used to identify patient populations having each disease of interest between January 1, 2018, and July 31, 2019. Patients could belong to multiple disease populations if they fulfilled multiple sets of criteria.

Data analysis
EHR data were collected and analyzed during September 2019. Each of the 8 disease populations was analyzed separately. For each disease population, a disease-specific problem list "status" was assigned to each patient based on the number of their outpatient problem list entries pertaining to that specific disease. Disease-specific entries were identified using the International Classification of Diseases–Tenth Revision (ICD-10) codes linked to each entry, as detailed in Supplementary Table 2. "Incomplete" problem lists had zero entries related to the disease. "Complete" problem lists had one or multiple entries related to the disease. "Duplicated" problem lists had multiple entries related to the disease. The "complete" and "duplicated" problem list statuses were both simultaneously assigned to any patient who had multiple entries related to the disease. Our definition of duplications assumed that ideally, only one problem list entry should exist for any single high-level ICD-10 code.

Additional data, such as age, sex, and lab results, were also collected. These additional parameters were used to assess disease severity in each patient, as detailed in Supplementary Table 3.

Using RStudio (v3.5.1) (Boston, MA), rates of problem list completeness and duplication for each disease population were calculated and compared according to disease type, patient demographics, and disease severity. In the age analyses, patients were separated into 3 age groups which roughly correspond to pediatric, adult, and geriatric populations: 0-20, 20-65, and 65+ years of age, respectively. Two-proportion z-tests were used to compare

Table 1. Baseline characteristics of each disease population

Disease	Mean age, y	Female (%)	Population Size	Problem list entries[a]
Asthma	52.4	62.1	38 502	1.10
Crohn's disease	45.1	52.0	3881	1.20
Depression	52.8	67.1	49 725	1.04
Diabetes	64.8	45.5	61 441	1.38
Epilepsy	48.2	50.7	11 292	1.13
Hypertension	66.9	52.5	233 535	0.78
Schizophrenia	49.0	34.3	995	0.92
Ulcerative colitis	49.1	54.6	3517	1.01

[a]Represents the average number of problem list entries related to the disease per patient.
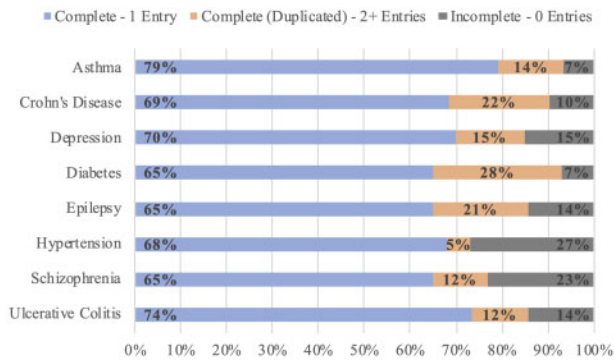

Figure 1. Rates of completeness, duplication, and incompleteness for each disease.

how rates of completeness and duplication varied according to patient demographics and disease severity. A *P* value of <.05 was considered statistically significant.

Regulatory compliance

This study was approved by the Partners HealthCare Institutional Review Board (Protocol #2009P001846) and the Harvard Medical School Institutional Review Board.

RESULTS

Overall population characteristics

A total of 327 695 unique patients and 383 404 problem list entries were identified across the 8 disease populations. For each disease population, baseline characteristics and average number of problem list entries related to the disease per patient are shown in Table 1.

Overall rates of completeness, duplication, and incompleteness

Figure 1 summarizes the overall rates of completeness, duplication, and incompleteness for the 8 diseases. Rates of completeness remained approximately in the 70%-90% range, with asthma having the highest rate at 93.5% and hypertension having the lowest rate at 72.9%. Rates of duplication varied more widely across the diseases. All diseases except hypertension had a duplication rate of at least 10%. Hypertension had the lowest rate of duplication at 4.8%, and the highest rate was 28.2% in diabetes. The proportion of patients with only one entry did not vary as much, remaining within the 65%-75% range for all diseases, except for asthma, which had a rate of 79.3%.
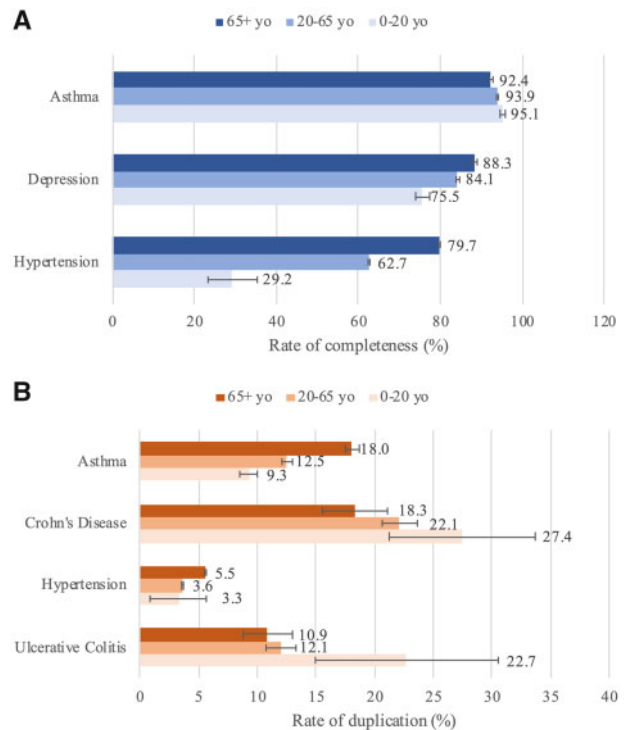

Figure 2. Relationship between age group and rates of completeness and duplications. Only diseases with a clear trend in this relationship are shown. We analyzed the relationship (A) between age group and rates of completeness and (B) between age group and rates of duplications.

Effect of demographics

For only certain disease populations, there was a relationship between demographic factors like age and sex, and rates of completeness and duplication. As shown in Figure 2, on the one hand, rates of completeness increased with increasing age in the depression and hypertension populations; on the other hand, completeness decreased with increasing age in patients with asthma. Rates of duplication also sometimes varied between different age groups. For patients with asthma and hypertension, rates of duplication increased with increasing age; on the other hand, rates of duplication decreased with increasing age in the Crohn's disease and ulcerative colitis populations. In the schizophrenia, diabetes, and epilepsy populations, there was no clear relationship between age group and rates of completeness and duplication (data not shown).

Similarly, rates of completeness and duplication varied according to patient sex for only some diseases, as summarized in Table 2. On the one hand, female patients with asthma, depression, or diabetes

Table 2. Relationship between patient sex and rates of completeness and duplication

Disease	Rates of completeness, %			Rates of duplication		
	Female, %	Male, %	Odds ratio	Female, %	Male, %	Odds ratio
Asthma	94.0 ± 0.3	92.7 ± 0.4	0.99[b]	14.7 ± 0.4	13.2 ± 0.5	0.90[b]
Crohn's disease	91.0 ± 1.2	89.5 ± 1.4	0.98	21.9 ± 1.8	21.3 ± 1.9	0.97
Depression	85.8 ± 0.4	83.4 ± 0.6	0.97[b]	15.1 ± 0.4	15.4 ± 0.6	1.02
Diabetes	93.6 ± 0.3	92.9 ± 0.3	0.99[b]	28.3 ± 0.5	28.2 ± 0.5	0.99
Epilepsy	85.6 ± 0.9	85.7 ± 0.9	1.00	21.0 ± 1.1	19.9 ± 1.1	0.95
Hypertension	72.0 ± 0.3	73.9 ± 0.3	1.03[b]	5.0 ± 0.1	4.6 ± 0.1	0.92[b]
Schizophrenia	73.9 ± 4.7	78.6 ± 3.1	1.06[a]	11.4 ± 3.4	12.5 ± 2.5	1.10
Ulcerative colitis	85.8 ± 1.6	85.5 ± 1.7	0.99	12.0 ± 1.5	12.3 ± 1.6	1.02

Values are mean ± SE. Odds ratios were calculated using the female group as the referent.
[a]$P < .05$; [b]$P < .001$.

Table 3. Relationship between disease severity and rates of completeness and duplication (all diseases, except diabetes and hypertension)

Disease	Rates of completeness			Rates of duplication		
	Less severe, %	More severe, %	Odds ratio	Less severe, %	More severe, %	Odds ratio
Asthma	93.4 ± 0.3	95.2 ± 1.3	1.02[a]	13.5 ± 0.3	38.4 ± 3.0	2.85[c]
Crohn's disease	88.7 ± 1.7	92.0 ± 1.3	1.04[c]	19.3 ± 2.1	27.0 ± 2.1	1.40[c]
Depression	84.4 ± 0.5	85.3 ± 0.4	1.01[b]	11.5 ± 0.5	17.0 ± 0.4	1.48[c]
Epilepsy	84.1 ± 0.7	96.3 ± 1.0	1.14[c]	16.1 ± 0.7	51.3 ± 2.6	3.20[c]
Schizophrenia	79.4 ± 4.9	76.1 ± 3.1	0.96	5.3 ± 2.7	14.6 ± 2.6	2.73[c]
Ulcerative colitis	83.1 ± 2.1	85.5 ± 2.0	1.03	11.4 ± 1.8	17.8 ± 2.2	1.56[c]

Values are mean ± SE. Odds ratios were calculated using the "less severe" disease group as the referent. Refer to Supplementary Table 3 for definitions of disease severity.
[a]$P < .05$; [b]$P < .01$; [c]$P < .001$.

had higher rates of completeness compared with male patients ($P < .001$); on the other hand, male patients with hypertension and schizophrenia had higher rates of completeness than female patients ($P < .05$). Rates of duplication were higher in female patients with asthma and hypertension than in male patients ($P < .001$). In Crohn's disease, epilepsy, and ulcerative colitis, there was no clear relationship between sex and rates of completeness and duplication.

Effect of disease severity

As shown in Table 3 and Figure 3, 6 of 8 diseases demonstrated a positive correlation between rates of completeness and disease severity, whereas all diseases demonstrated a strong positive correlation between rates of duplication and disease severity.

For Crohn's disease, patients with more severe disease had a higher rate of completeness at 92.0%, compared with 88.7% in their counterparts with less severe disease ($P < .001$). In the ulcerative colitis population, patients with more severe disease had a slightly higher rate of completeness at 85.5%, compared with 83.1% in their counterparts, but this was not statistically significant. For both Crohn's disease and ulcerative colitis, patients with more severe disease had higher rates of duplication at 27.0% and 17.8%, respectively, compared with 19.3% and 11.4% in their counterparts ($P < .001$).

For patients with depression, the rate of completeness was higher at 85.3% in those with more severe depression, compared with 84.4% in their counterparts ($P < .01$). In patients with schizophrenia, there was no significant difference in rates of completeness between the 2 groups. For both disease populations, rates of

duplication were higher in patients with more severe disease. Patients with more severe depression had a higher duplication rate of 17.0%, compared with 11.5% in their counterparts ($P < .001$), and patients with more severe schizophrenia had a higher duplication rate of 14.6%, compared with 5.3% in their counterparts ($P < .001$).

In patients with diabetes, rates of completeness and duplication were generally positively correlated with average hemoglobin A1c (HgbA1c) score, as shown in Figures 3A and 3B. Patients with HgbA1c scores in the 6.5%-7.5% range had the second-lowest rate of completeness at 92.6%, as shown in Figure 3A. This was significantly lower than the rates of completeness for patients in the 7.5%-8.5%, 8.5%-9.5%, and 9.5%-10.5% ranges ($P < .02$). Among these latter 3 groups, there was no significant difference in rates of completeness. Interestingly, patients with HgbA1c scores above 10.5% had the lowest rate of completeness of all 5 groups, at 91.1% ($P < .001$). The results for rates of duplication were similar, as shown in Figure 3B. Patients with HgbA1c scores in the 6.5%-7.5% range had the lowest rate of duplication, at 24.4% ($P < .001$). The rate of duplication increased to 29.7%, 33.9%, and 34.8%, respectively, for the next 3 consecutive HgbA1c score groups. For patients with scores above 10.5%, the rate of duplication was 32.4%, which was lower than that of the 9.5%-10.5% HgbA1c group directly below ($P < .02$).

Similar findings were observed in patients with hypertension, as shown in Figures 3C and 3D. In patients with hypertension, rates of completeness generally increased with increasing average systolic blood pressure (SBP), as shown in Figure 3C. The rate of completeness was lowest at 47.6% for patients with an average SBP of
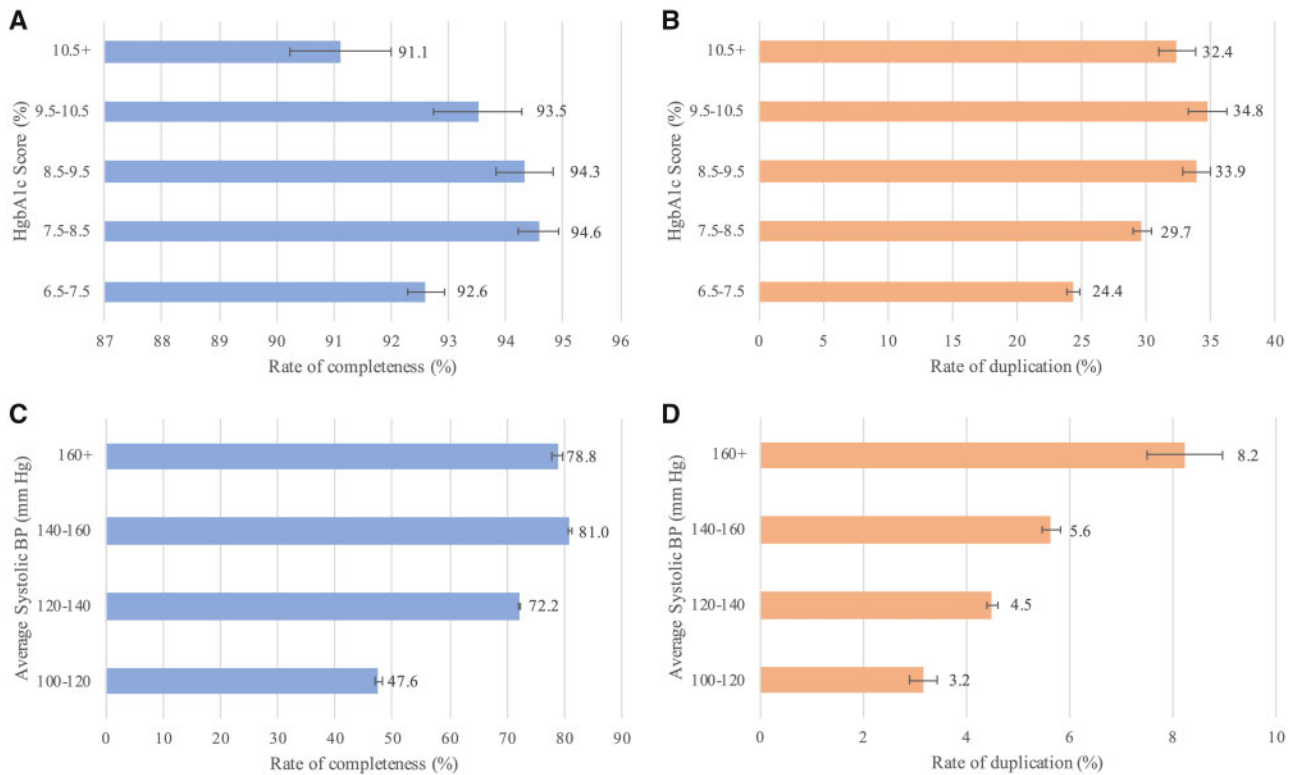
Figure 3. Relationship between disease severity and rates of completeness and duplications for diabetes and hypertension. We analyzed the relationship between average hemoglobin A1c (HgbA1c) score and rates of (A) completeness and (B) duplications. We analyzed the relationship between average systolic blood pressure and rates of (C) completeness and (D) duplication.

100-120 mm Hg, and it was highest at 81.0% in patients with an average SBP of 140-160 mm Hg. Similar to the findings in the diabetes population, the rate of completeness in the group with the highest SBP (>160 mm Hg) was the second highest, at 78.8%. Rates of duplication were strongly positively correlated with average SBP, as shown in Figure 3D. The 95% confidence intervals for all 4 groups did not overlap in the analysis of duplications.

In patients with asthma, the rate of completeness for patients with more severe asthma was higher at 95.2%, compared with 93.4% in those with less severe asthma (*P* < .02). Similarly, those with more severe asthma had higher rates of duplication at 38.4%, compared with 13.5% in their counterparts (*P* < .001).

Similar trends were observed in patients with epilepsy. Rates of completeness were higher in those with more severe epilepsy, at 96.3%, compared with 84.1% in their counterparts (*P* < .001). Rates of duplication were also higher in those with more severe epilepsy, at 51.3%, compared with 16.1% in their counterparts (*P* < .001).

DISCUSSION

Overview of rates of completeness, duplication, and incompleteness

Rates of completeness were relatively stable in the 70%-90% range across all diseases. Rates of duplication were generally greater than rates of incompleteness, averaging 16.1% and 14.6% across the diseases, respectively, suggesting that the burden of duplications was slightly higher than that of incompleteness in these populations.

Across the diseases, the proportion of patients specifically with only 1 problem list entry remained relatively consistent in the 65%-75% range, except for asthma. In contrast, the proportion of patients with duplicated entries was much more variable, ranging from 4.8% in hypertension to 28.2% in diabetes. This suggests that most of the variation in rates of completeness is specifically attributable to variations in rates of duplication, not the rates of single-entry problem lists. In other words, providers consistently provide 65%-75% of patients with 1 and only 1 problem list entry, regardless of the disease. However, for some diseases more than others, providers have a variable tendency to create duplications, likely because of disease-specific differences in the number of problem list–encodable attributes, such as complications and disease stages.

Effect of demographics

In only some diseases, we found a relationship between demographics and rates of completeness and duplications. This relationship was most apparent in diseases that are more prevalent within a specific demographic group. In hypertension, which is more prevalent in older adults, rates of completeness were positively correlated with age, whereas the opposite trend was observed in asthma, likely because asthma is more common in children.

There were similar findings with regard to sex. Depression and hypertension, which are more prevalent among female and male patients, respectively, demonstrated higher rates of completeness in these same demographic groups. Of note, this relationship was reversed when evaluating rates of duplication in patients with hypertension, with female patients having a higher rate of duplication than male patients. One possible explanation is that some

hypertension-related problem list entries are more specific to female patients, such as "gestational hypertension."

All of these findings suggest that there is no universal relationship between demographics and rates of completeness and duplication. Even when there is a relationship, its effect size is small, especially when compared with that of disease severity. This effect also tends to be smaller for completeness than for duplications. Nevertheless, a relationship is more likely to be present if a disease is more prevalent in a specific, "classic" demographic group. If a patient matches the "classic" presentation, they may be more likely to have a complete problem list—either because the provider is more primed to remember to maintain a complete problem list, or because the provider is more confident about officially listing the diagnosis in the problem list. For patients with atypical presentations, the opposite is likely true, and they may be more likely to have an incomplete problem list and all of its associated detrimental effects. Thus, future EHR interventions might focus on providing additional prompts for providers who are treating patients with less typical presentations in order to avoid perpetuating this bias.

Effect of disease severity

Compared with demographics, there was a more consistent relationship between disease severity and completeness and duplications, with only 2 diseases failing to demonstrate a relationship with completeness. The effect of disease severity on completeness was similar in magnitude to that of demographics, but the effect of disease severity on duplications was substantially greater than that for both completeness and demographics. One notable finding is that in the diabetes population, patients with HgbA1c scores above 10.5% had lower rates of completeness and duplication than was expected. This finding was also observed in patients with hypertension who had average SBP above 160 mm Hg. One possible explanation for these unexpected findings is that patients who have the highest HgbA1c scores or average SBP represent patients with less access to care or those with more uncontrolled, but asymptomatic, disease who may not have been diagnosed until only recently; therefore, providers may have had fewer encounters with these patients and had fewer opportunities to create complete or duplicated problem lists.

The positive correlation between duplications and disease severity may be the result of more severe disease states leading to more systemic complications that can be separately documented in the problem list. Patients with more severe disease may also have more healthcare providers and more encounters with the healthcare system, leading to a concomitant increase in duplications and documentation in general. The problem may then be exacerbated by providers' unwillingness to edit or remove other providers' entries, leading to the inevitable accumulation of duplications. Finally, problem list entries in this particular EHR are often automatically copied from billing diagnoses, which financially incentivize the explicit listing of all complications, thereby exacerbating the issue.

In contrast to the consistent positive correlation between duplications and disease severity, there was a weaker, less consistent positive correlation between completeness and disease severity. This suggests that providers are relatively better at maintaining complete problem lists for all patients, regardless of disease severity. However, they struggle more with maintaining complete problem lists for patients with less severe and less symptomatic disease, likely because these patients have fewer providers and fewer visits—both of which limit the opportunities to maintain the problem list. Future EHR interventions should continue to address the issue of incompleteness, especially for patients with insidious or asymptomatic disease, which would otherwise have a higher likelihood of being underrepresented in the problem list. Future interventions should also reduce duplications, especially in patient populations with high healthcare utilization and more severe disease because these populations are more vulnerable to the detrimental effects of duplications. Even simply organizing duplications by disease or system may be immensely effective in reducing their visual and cognitive burden.

Limitations

Our dataset was limited to EHR data from Partners HealthCare, which predominantly provides tertiary and quaternary academic referral care in a metropolitan area. As a result, the patients and providers in this healthcare organization are not necessarily representative of the general U.S. population. Our EHR dataset was derived from the Epic commercial EHR only and did not capture any healthcare information from other healthcare organizations. The study was also limited in scope to 8 chronic diseases.

There are also inherent limitations to using EHR data. Many patient records had incomplete data, such as missing blood pressure measurements and lab values. For the purpose of calculations and statistical tests, any missing values were censored, which likely introduces bias into our analyses because patients with missing data may be inherently different from those with all of the available data. Furthermore, the definitions for each of our diseases were limited to the structured data within the EHR database and therefore did not incorporate clinical input from healthcare providers. As a result, our method of identifying patients with each disease may have failed to capture patients with well-controlled disease, patients without the requisite lab values, and patients who have a significant amount of care provided by other organizations. Similarly, our definitions of problem list completeness and duplications were solely dependent on the number of problem list entries, rather than relying on the actual clinical information contained within the entries themselves. Our definition of duplications also assumed that there should ideally be no more than one entry for any high-level ICD-10 code; this assumption was likely oversensitive because it assumed that all duplicates were not useful, which is not always the case. We are not aware of any prior literature that defines duplications differently. We also relied on ICD-10 codes to determine which problem list entries belonged to which disease. Without manual review of each chart, this method failed to capture any entries that were created manually and also precluded us from analyzing the semantic information within the comments section of each problem list entry.

Finally, for all diseases except depression, asthma, and epilepsy, there were significantly smaller sample sizes within the pediatric populations, leading to the potential for more random variation in this specific population.

Suggestions for future work

Because our study focused on characterizing aggregate rates of completeness and duplication strictly at the disease-level, future studies should assess how rates of completeness and duplication vary by provider or healthcare setting. It is likely that individual provider usage habits or healthcare organization policies play a large role in shaping the overall patterns of problem list utilization. Similarly, software design may affect the way problem lists are utilized, so repeating this study with another EHR vendor may potentially yield different results. Our study also did not evaluate the unstructured

content of the comments section in problem lists. Future studies into the unstructured content of problem lists are warranted and may provide insight into the types of unstructured information that providers often input into problem lists.

While our findings provide novel, more thorough characterizations of the current shortcomings in problem list utilization patterns, they do not elucidate why these shortcomings occur in the first place or how these shortcomings affect efficiency, quality, and safety. This is a domain that we plan to study in the future. Future studies should qualitatively analyze provider behaviors, usage patterns, and opinions with regard to problem list utilization, completeness, duplication, and maintenance. This information will be instrumental to the future development of organizational policies and EHR interventions that can address these issues. Future interventions should directly reduce the burden of problem lists currently being shouldered by providers. This might involve natural language processing and other tools that automatically manage problem lists based on data, especially notes, that providers already generate. Institutions and professional organizations also play a role in promoting best practices and policies that improve problem list utilization.[12] Last, prudent regulation can encourage effective problem list utilization by linking best practices to financial incentives and quality measurements.

CONCLUSION

We analyzed problem list utilization across 8 chronic diseases, focusing on how problem list completeness and duplications vary according to disease type, demographics, and disease severity. Overall, rates of problem list completeness and duplication varied widely across the different diseases. Demographic factors such as age and sex did not consistently affect problem list completeness and duplication, except in certain cases in which demographics are essential to the medical conception of the disease. In contrast, we found that problem list completeness and duplications are often positively correlated with disease severity. These findings demonstrate the shortcomings of current problem list utilization habits and highlight the need for further research into this vital function of the EHR.

FUNDING

AUTHOR CONTRIBUTIONS

EC-HW formulated the research question and conducted all database queries and analyses. AW provided valuable mentorship, expertise, and guidance throughout the entire process. EC-HW led the writing of this article, and AW provided comments and edits on all drafts.

ACKNOWLEDGMENTS

SUPPLEMENTARY MATERIAL

Supplementary material is available at *Journal of the American Medical Informatics Association* online.

CONFLICT OF INTEREST STATEMENT

None declared.

REFERENCES

1. Holmgren AJ, Co Z, Newmark L, *et al*. Assessing the safety of electronic health records: a national longitudinal study of medication-related decision support. *BMJ Qual Saf* 2020; 29 (1): 52–9.
2. McGinn TG, McCullagh L, Kannry J, *et al*. Efficacy of an evidence-based clinical decision support in primary care practices: a randomized clinical trial. *JAMA Intern Med* 2013; 173 (17): 1584–91.
3. Payne TH, Corley S, Cullen TA, *et al*. Report of the AMIA EHR-2020 Task Force on the status and future direction of EHRs. *J Am Med Inform Assoc* 2015; 22 (5): 1102–10.
4. Evans RS. Electronic health records: then, now, and in the future. *Yearb Med Inform* 2016; 25 (S01): S48–61.
5. Grabenbauer L, Skinner A, Windle J. Electronic health record adoption: maybe it's not about the money. *Appl Clin Inform* 2011; 2 (4): 460–71.
6. Downing NL, Bates DW, Longhurst CA. Physician burnout in the electronic health record era: are we ignoring the real cause? *Ann Intern Med* 2018; 169 (1): 50–1.
7. Gardner RL, Cooper E, Haskell J, *et al*. Physician stress and burnout: the impact of health information technology. *J Am Med Inform Assoc* 2019; 26 (2): 106–14.
8. Moros DA. The electronic medical record and the loss of narrative. *Camb Q Healthc Ethics* 2017; 26 (2): 328–31.
9. Simborg DW, Starfield BH, Horn SD, *et al*. Information factors affecting problem follow-up in ambulatory care. *Med Care* 1976; 14 (10): 848–56.
10. Joseph JW, Chiu DT, Nathanson LA, *et al*. A rules based algorithm to generate problem lists using emergency department medication reconciliation. *Int J Med Inform* 2016; 94: 117–22.
11. Hartung DM, Hunt J, Siemienczuk J, *et al*. Clinical implications of an accurate problem list on heart failure treatment. *J Gen Intern Med* 2005; 20 (2): 143–7.
12. Wright A, Maloney FL, Feblowitz JC. Clinician attitudes toward and use of electronic problem lists: a thematic analysis. *BMC Med Inform Decis Mak* 2011; 11 (1): 36.
13. Kaplan DM. Clear writing, clear thinking and the disappearing art of the problem list. *J Hosp Med* 2007; 2 (4): 199–202.
14. Szeto HC, Coleman RK, Gholami P, *et al*. Accuracy of computerized outpatient diagnoses in a Veterans Affairs general medicine clinic. *Am J Manag Care* 2002; 8 (1): 37–43.
15. Van Vleck TT, Wilcox A, Stetson PD, *et al*. Content and structure of clinical problem lists: a corpus analysis. *AMIA Annu Symp Proc* 2008; 2008: 753–7.
16. Wright A, McCoy AB, Hickman TT, *et al*. Problem list completeness in electronic health records: A multi-site study and assessment of success factors. *Int J Med Inform* 2015; 84 (10): 784–90.
17. Holmes C, Brown M, Hilaire DS, *et al*. Healthcare provider attitudes towards the problem list in an electronic health record: a mixed-methods qualitative study. *BMC Med Inform Decis Mak* 2012; 12 (1): 127.
18. Wright A, Feblowitz J, Maloney FL, *et al*. Use of an electronic problem list by primary care providers and specialists. *J Gen Intern Med* 2012; 27 (8): 968–73.
19. Devarakonda MV, Mehta N, Tsou CH, *et al*. Automated problem list generation and physicians perspective from a pilot study. *Int J Med Inform* 2017; 105: 121–9.
20. Wright A, Pang J, Feblowitz JC, *et al*. A method and knowledge base for automated inference of patient problems from structured data in an electronic medical record. *J Am Med Inform Assoc* 2011; 18 (6): 859–67.
21. Wright A, Pang J, Feblowitz JC, *et al*. Improving completeness of electronic problem lists through clinical decision support: a randomized, controlled trial. *J Am Med Inform Assoc* 2012; 19 (4): 555–61.
22. Vogel L. Cut-and-paste clinical notes confuse care, say US internists. *CMAJ* 2013; 185: e826.

23. Hammond KW, Helbig ST, Benson CC, *et al*. Are electronic medical records trustworthy? Observations on copying, pasting and duplication. *AMIA Annu Symp Proc 2003;* 2003: 269–73.

24. Siegler EL, Adelman R. Copy and paste: a remediable hazard of electronic health records. *Am J Med* 2009; 122 (6): 495–6.

25. Khurana S, Jarjour NN. Systematic approach to asthma of varying severity. *Clin Chest Med* 2019; 40 (1): 59–70.

26. Flynn S, Eisenstein S. Inflammatory bowel disease presentation and diagnosis. *Surg Clin N Am* 2019; 99 (6): 1051–62.

27. Malhi GS, Mann JJ. Depression. *Lancet* 2018; 392 (10161): 2299–312.

28. Sacks DB, John WG. Interpretation of hemoglobin A1c values. *JAMA* 2014; 311 (22): 2271–2.

29. Manjunath R, Paradis PE, Parisé H, *et al*. Burden of uncontrolled epilepsy in patients requiring an emergency room visit or hospitalization. *Neurology* 2012; 79 (18): 1908–16.

30. Bakris G, Sorrentino M. Redefining hypertension: assessing the new blood-pressure guidelines. *N Engl J Med* 2018; 378 (6): 497–9.

31. Schuitz SK, Andreasen NC. Schizophrenia. *Lancet* 1999; 353 (9162): 1425–30.