Data Gaps in Electronic Health Record (EHR) Systems: An Audit of Problem List Completeness During the COVID-19 Pandemic

Supplementary data

Jordan Poulos, Leilei Zhu, Anoop D. Shah

Table of Contents

1. Further information on methods to improve problem list completeness

Our paper references several studies which have identified several methods to improve problem list uptake and completeness. This section provides further information on these references.

1) Automatically populating problem lists from diagnosis information recorded elsewhere in the EHR. At the Hospital Italiano de Buenos Aires (HIBA), automated loading of patient co-morbidities from a separate structured data section of the EHR into problem lists with clinician approval resulted in a 386% increase in the number of problem list terms appearing on problem lists [1]. Problem-based charting (PBC), when integrated with the electronic problem list, increased recording of sepsis and other common medical conditions in intensive care unit records at Stamford Hospital [2].

2) Educational interventions combined with integration of problem lists into various aspects of EHR functionality. A quality improvement project at the Department of Paediatric Hospital Medicine at Cleveland Clinic Children's demonstrated that educational interventions to acknowledge the relevance and significance of appropriate problem list use, in combination with technological leveraging of EHR resources (adding problem lists to discharge navigators and admission order sets, and encouraging problem based progress notes) resulted in 100% of all inpatient charts with at least one documented problem in the problem list, from a baseline of 47% [3]. Clinician alters to add to the problem list when undertaking other activities in the EHR, such as prescribing, have also been shown to be effective [4].

3) Increased patient access and involvement in regulating the problem list. Self-reporting of chronic conditions has demonstrated high specificity [5-7], and hold the potential to verify existing problem list entries and identify possible gaps in the problem list [8].

4) Improving the accuracy of problem lists through machine learning and natural language processing (NLP). NLP techniques to automatically update problem lists from electronic notes are being refined, but already have shown to demonstrate good sensitivity compared to clinician gold standard [9-12].

2. Additional background information

In this appendix, the authors provide further background information as to why the study was undertaken, and why the cohort of suspected and confirmed COVID-19 inpatients was selected.

The 2020 COVID-19 pandemic highlighted the need for high quality EHR data for rapid service planning, operational reporting, and clinical research. The pandemic occurred one year following the installation of Epic, a comprehensive EHR system at UCLH (University College London Hospitals) Trust (in April 2019). All inpatient and outpatient clinical notes, observations, prescriptions, drug charts and order communications are now digital.

The EHR system was rapidly adapted to add a bespoke data item for COVID-19 status, which was displayed as an infection flag on the patient banner and inpatient lists. The infection flag was set by the infectious diseases team based on laboratory results and the patient's clinical status, and enabled the cohort of patients with COVID-19 to be reliably identified. However, other data items collected by the registries were less consistently recorded in the structured data fields of the EHR, and instead were available only in free text electronic clinical notes.

UCLH made funding available for departments to employ medical students to support the pandemic effort. The EHR Directorate therefore recruited a group of medical students to assist in correcting the deficiencies in the availability of structured EHR data in the problem list (diagnoses) and a number of other key data fields (ethnicity, smoking history, clinical frailty score, and date of symptom onset for COVID-19 patients) under clinical supervision. This would improve the ease of retrieving key EHR data for future patient encounters, and improve the quality of EHR-derived research datasets such as the National Institute for Health Research (NIHR) Health Informatics Collaborative (https://hic.nihr.ac.uk/covid) and the DECOVID collaboration with University Hospitals Birmingham and the Turing Institute (https://www.turing.ac.uk/research/research-projects/decovid). It would also enable easier extraction of data for COVID-19 research registries such as ISARIC and CAPACITY-COVID, reducing the need for duplicate data entry. Comprehensive retrospective chart review of this specific cohort of patients was also required to support research at the trust to evaluate prognostic models for COVID-19 [13].

3. Standard Operating Procedure

The following document is the standard operating procedure (SOP) used to standardise the process of repopulating the problem list from free-text entries in the electronic notes.

Standard operating procedure for updating problem list for patients with suspected or confirmed COVID-19

The problem list is a core component of the EHR system. The problem list combines past medical history, current diagnoses and other important health issues. They are classified as 'Active' problems, which are those requiring current medical attention, and 'Resolved' problems, which are inactive or past medical history.

1. Run report to generate a list of patients to review (e.g current patients at the trust with active status of COVID-19). Carry out the following steps for each patient:
2. Use notes section to find clerking notes for patient past medical history (PMH). Double check no PMH has been missed by using the chart search function and search ‘PMH’. Compare with PMH notes only from the current admission. For current additional problems check the most recent ward round notes/timeline for ongoing hospital issues.
3. Add each PMH to the problem list, using the terms explicitly stated in the clerking notes.
4. Leave the date noted as the current date. If there is a specific date of onset for the problem, this should be written in the overview section as free text. Select ‘accept’ to save these changes.
5. Move all PMH problems which are neither the cause of admission, nor have been diagnosed during the admission, by unticking the ‘Hospital’ box. If there is any uncertainty of a PMH problem, refer to a clinician for advice.
6. To add the COVID-19 diagnosis, select the patient infection status toolbar and note the first date the infection flag was added by an infection control nurse. This information can be added as the ‘noted’ date.
7. Mark problem list as reviewed.

4. Cohort demographics

Aggregated table of patient demographics in the cohort (n=516)

Sex	**Number of patients**
Male	336
Female	180

Ethnicity	**Number of patients**
White	290
Black	72
South Asian	73
Mixed or Other	58
Missing	23

Age	**Number of patients**
<20	5
20-29	15
30-39	31
40-49	56
50-59	80
60-69	131
70-79	86
80-89	87
90-99	24

References

1. Fosser MS, Gaiera MA, Otero MC, et al. Automatic loading of problems using a COMORBIDITIES subset, one step to organize and maintain the patient's problem list. *Annals of Family Medicine.* 2006;4:101-3.10.
2. Li RC, Garg T, Cun T, et al. Impact of problem-based charting on the utilization and accuracy of the electronic problem list. *Journal of the American Medical Informatics Association.* 2018 May;25(5):548-54.
3. Rajbhandari P, Auron M, Worley S, et al. Improving Documentation of Inpatient Problem List in Electronic Health Record: A Quality Improvement Project. *Journal of patient safety.* 2018 Apr 19.
4. Galanter WL, Hier DB, Jao C, et al. Computerized physician order entry of medications and clinical decision support can improve problem list documentation compliance. *International journal of medical informatics*. 2010 May 1;79(5):332-8.
5. Natarajan S, Lipsitz SR, Nietert PJ. Self-report of high cholesterol: determinants of validity in US adults. *American journal of preventive medicine*. 2002 Jul 1;23(1):13-21.
6. Goldman N, Lin IF, Weinstein M, et al. Evaluating the quality of self-reports of hypertension and diabetes. *Journal of clinical epidemiology.* 2003 Feb 1;56(2):148-54.
7. Goto A, Morita A, Goto M, et al. Validity of diabetes self-reports in the Saku diabetes study. *Journal of epidemiology*. 2013 Jul 5:JE20120221.
8. Fernanda Polubriaginof CG, Paulo Pastore G, Drohan B, et al. Comparing Patient-Reported Medical Problems with the Electronic Health Record Problem List. *Gen Med (Los Angel)*. 2016;4(258):2.
9. Meystre S, Haug P. Improving the sensitivity of the problem list in an intensive care unit by using natural language processing. InAMIA Annual Symposium Proceedings 2006 (Vol. 2006, p. 554). *American Medical Informatics Association.*
10. Meystre S, Haug PJ. Automation of a problem list using natural language processing. *BMC medical informatics and decision making.* 2005 Dec;5(1):1-4.
11. Plazzotta F, Otero C, Luna D, et al. Natural language processing and inference rules as strategies for updating problem list in an electronic health record. *Studies in health technology and informatics.* 2013 Jan 1;192:1163-.
12. Devarakonda MV, Mehta N, Tsou CH, et al. Automated problem list generation and physicians perspective from a pilot study. *International journal of medical informatics.* 2017 Sep 1;105:121-9.
13. Gupta RK, Marks M, Samuels TH, et al. Systematic evaluation and external validation of 22 prognostic models among hospitalised adults with COVID-19: an observational cohort study. *European Respiratory Journal.* 2020 Dec 1;56(6).