MEDINFO 2015: eHealth-enabled Health
I.N. Sarkar et al. (Eds.)

doi:10.3233/978-1-61499-564-7-158

Characterizing Patient-Generated Clinical Data and Associated Implications for Electronic Health Records

Elliot G. Arsoniadis[a, b], Rabindra Tambyraja[c], Saif Khairat[d], Cyrus Jahansouz[b], Daren Scheppmann[e], Mary R. Kwaan[b], Gretchen Hultman[a], Genevieve B. Melton[a, b]

[a]*Institute for Health Informatics, University of Minnesota, Minneapolis, MN, USA*
[b]*Department of Surgery, University of Minnesota, Minneapolis, MN, USA*
[c]*Department of Psychiatry, University of Minnesota, Minneapolis, MN, USA*
[d]*Carolina Health Informatics Program, University of North Carolina, Chapel Hill, NC, USA*
[e]*Fairview Health Services, Minneapolis, MN, USA*

Abstract

Patient-facing technologies are increasingly utilized for direct patient data entry for potential incorporation into the electronic health record. We analyzed patient-entered data during implementation of a patient-facing data entry technology using an online patient portal and clinic-based tablet computers at a University-based tertiary medical center clinic, including entries for past medical history, past surgical history, and social history. Entries were assessed for granularity, clinical accuracy, and the addition of novel information into the record. We found that over half of patient-generated diagnoses were duplicates of lesser or equal granularity compared to previous provider-entered diagnoses. Approximately one fifth of patient-generated diagnoses were found to meet the criteria for new, meaningful additions to the medical record. Our findings demonstrate that while patient-generated data provides important additional information, it may also present challenges including generating inaccurate or less granular information.

Keywords:

Patient generated healthcare data; Patient facing technology; Electronic Health Record;

Introduction

As patient-centeredness is an increasingly important tenet in healthcare, technologies engaging patients for individual assessment of clinical status, satisfaction, and education are being implemented in larger number. Moreover, the importance of patient engagement for attracting and maintaining patients is increasingly recognized by health systems [1]. As part of this, the electronic health record (EHR) has undergone significant change, including the integration of patient-facing portals with the EHR. Increasingly, portals and other patient-facing tools allow patients to access their own medical data and in some cases grant the ability to modify or comment on the data in these platforms [2].

A number of studies have shown that patients find the use of a patient portal for direct data entry to be a positive experience. Wu et al. reported 98% of patients in their large cohort found the use of a clinic-based kiosk for patient-generated family health history information to be useful [3]. A similar study found over 90% satisfaction with a similar online patient-facing family health history tool in a population of Veterans Administration patients [4]. Patient-facing tools relying on patient-generated data for realtime feedback have also been used with success in the areas of lifestyle modification [5] and chronic disease management [6].

To date, there are few studies assessing the quality of patient-generated healthcare data. While there have been some reports of positive effects with the addition of clinically relevant data [7,8], these studies have been limited to certain diagnoses, such as cardiac disease, or to certain portions of the medical record, such as family history [4,8]. In this study, we sought to better understand patient entered data broadly for different aspects of health history through an evaluation of the quality of patient-generated data using a patient-facing health history tool for past medical history (PMH), past surgical history (PSH), and social history (SH).

Materials and Methods

The patient-facing tool evaluated in this study was available September 2014 to patients with appointments at a surgery clinic for the University of Minnesota Physicians. The content of the tool was an organization-wide health history questionnaire approved by the medical directors of the practice plan. The tool was available to patients enrolled in the EHR online patient portal prior to the clinic appointment. For those patients not enrolled in the online patient portal or with portal access but who failed to utilize the tool prior to presentation in clinic, tablet computers were made available on site for completion prior to meeting with the provider. Previous studies have shown that patients respond well to tablet computer-based health applications, despite bias that these technologies might be unusable in certain patient populations, such as the elderly [9].

The questionnaires were only available in English and covered PMH, PSH, and SH. Patients were presented with a list of 40 common medical diagnoses, 22 surgical procedures, and four social history domains. PMH and PSH sections included "Yes" or "No" options, as well as a "Comments" section for free-text of additional conditions and procedures. The SH portion of the tool included questions related to tobacco use, alcohol consumption, illicit substance use, and sexual history. For the first three, patients were asked about use, type, amount, duration, and date of cessation (if applicable). The sexual history domain included current status, partner gender, and any birth control measures.

Following completion of the tool, results were immediately available to the clinician on the enterprise Epic EHR system. The provider then was presented with the opportunity to "accept" or "reject" patient entered information. If accepted, the entries were then associated with discrete diagnosis terms and diagnosis codes in the PMH and PSH sections, as well as the discrete data fields of the SH section.

Table 1 – Physician-rater Diagnosis Coding

Physician-rater grade	Example
Duplicate Diagnosis Diagnosis already in PMH/PSH. Granularity of diagnosis is equal or near equal to that in the PMH/PSH.	Patient enters "Hypertension". "Hypertension" in PMH.
Duplicate Low Granularity Diagnosis Diagnosis already in PMH/PSH. Granularity of diagnosis is less that that in the PMH/PSH.	Patient enters "Kidney Problems". "Diabetic nephropathy" in PMH - Patient enters "Fracture or Broken Bone". "L3 compression fracture" in PMH.
New Low Granularity Diagnosis Diagnosis not in PMH/PSH and insufficient granularity to provide clinical utility.	Patient enters "Respiratory Problem". No previous respiratory diagnosis listed. Patient enters "Fracture or broken bone" but fails to mention anatomic location.
New False Diagnosis Diagnosis not in PMH/PSH and with sufficient granularity to provide clinical utility but on EHR review appears unlikely to be true.	Patient lists "Blood Clotting Disorder". Clinician notes "no evidence of thrombophilia" in Hematology clinic progress note.
New Diagnosis with Utility Diagnosis not in PMH/PSH and with sufficient granularity to provide clinical utility and corroborating EHR information.	Patient enters diagnosis of "Anxiety". No mental health problem is listed in PMH. Review of patient's medication list shows benzodiazepine. Patient enters a history of cholecystectomy. None is listed in PSH. CT scan report reads "gallbladder surgically absent".

We reviewed the data of 50 patients who completed the tool within the first six weeks of implementation, out of a possible 146 eligible participants (response rate 34.2%). Inclusion criteria included any new or return patient to the clinic who previously had been seen within the institution. A "New" visit type was designated for a visit where the patient had not been seen in surgery clinic before and were new to the provider, but had been previously seen by another provider (outside of our surgery clinic) within the Fairview Healthsystem. Since our analysis included comparing patient-generated data against data already contained within the enterprise EHR, patients new both to the clinic and to the healthcare system were excluded from this study. All patient-generated diagnoses were reviewed, regardless of acceptance by the provider.

As summarized in Table 1, each diagnosis entered was then reviewed by an independent physician-rater. The initial step required the physician to determine whether the patient-generated diagnosis was previously listed in the PMH and PSH of the patient chart. The next step involved gauging the level of granularity provided by the patient-generated diagnosis. For those diagnoses previously listed in the EHR, the rater then determined if the patient-generated diagnosis was either equal in granularity or lesser in granularity compared to the initial diagnosis. The physician-rater also gauged the granularity of any diagnosis deemed "New" on review in the initial step. Patient-generated diagnoses lacking sufficient granularity to enrich the clinical picture of the patient were classified as "New Low Granularity Diagnosis". Those patient-generated diagnoses judged to be of sufficient granularity to provide enrichment of the clinical history underwent a third level of scrutiny. In this final step, the physician-rater reviewed the patient chart to look for evidence to suggest if the patient-generated diagnosis was likely to be accurate. Clinical narratives from physicians, both from within and outside of the health system, were utilized as well as imaging studies, relevant laboratory values, medication lists, and procedure/endoscopy notes, as appropriate. Inter-rater reliability was assessed in five (ten percent) patient charts by a second physician-rater. Percent agreement and Kappa were calculated for coding of PMH/PSH diagnoses/procedures. Institutional review board approval was obtained and informed consent waived for this minimal risk study.

Results

Information about the patients and their types of visits are summarized in Table 2. In total, the 50 patients had a total of 435 patient-generated medical diagnoses, 231 surgical procedures, and 188 social history elements associated with these visits.

Table 2 – Patient Demographics and Visit Type

Demographic	n (%)
Age, range (mean)	18-74 (49.0)
Gender (female)	26 (52.0)
Visit Type	
New	6 (12.0)
Return	44 (88.0)
History Information	
Patient-generated diagnoses	435
Patient-generated procedures	231
Patient-generated social history	188

Over 50 percent of patient-generated medical diagnoses were duplicates of diagnoses already contained within the PMH portion of the EHR (Table 3). Of these duplicates, slightly greater than half of patient-generated diagnoses were of equal granularity to previous provider-entered diagnoses. The remaining half of the duplicated diagnoses had less granularity compared to previous provider-entered diagnoses. Nearly ten percent of the patient-generated diagnoses were determined to be new, but lacked sufficient granularity to provide meaningful clinical enrichment (New Low Granularity Diagnosis). Of the patient-entered diagnoses determined to be both new and providing sufficient granularity to be considered clinically meaningful, nearly half were corroborated on chart review with the other half lacking sufficient evidence within the EHR to verify diagnosis (New Diagnosis with Utility and New False Diagnosis, respectively). Over 20 percent of new diagnoses deemed true based on chart review were related to mental health conditions (13.8% anxiety, 7.5% depression). History of colon/rectal polyps (11.3%) and history of blood transfusion (8.8%) were also common patient-generated diagnoses that were found to be true. Inter-rater reliability with a second physician-rater resulted in percent correlation of 0.89 and Kappa statistic of 0.817.

Forty-six of the 50 patients in this study completed the surgical history section of the survey. Two patients listed no previous surgeries. A total of 231 surgical procedures were entered by the remaining 44 patients (Table 4). Nearly 75% of surgical procedures were duplicates of procedures already recorded within the PSH portion of the EHR. Twenty-three percent were determined by an independent physician-rater to provide equal granularity to procedures previously listed. Nearly 52% were found to be of insufficient granularity compared to the surgical history already available in the EHR. The remaining surgical procedures entered were new to the PSH portion of the EHR. Of these, nearly half (ten percent of total) were determined to provide insufficient granularity to be of clinical utility. However, 11.3% of the total surgical procedures entered by patients were found to be new, sufficiently granular, and had evidence within the EHR. Three percent of surgical procedures entered by patients were new and sufficiently granular, but had no evidence within the EHR. One patient entered the same surgical procedure twice (Duplicate Entry – 0.9% of total). Inter-rater reliability with a second physician-rater resulted in percent correlation of 0.9 and Kappa statistic of 0.837.

Table 3 – Past Medical History

Category	n (%)
Duplicate Diagnosis	117 (26.9)
Duplicate Low Granularity Diagnosis	112 (25.7)
New Low Granularity Diagnosis	43 (9.9)
New False Diagnosis	82 (18.9)
New Diagnosis with Utility	80 (18.4)

All 50 patients completed at least one of the four domains within the SH portion of the survey. Overall, 188 data elements relating to tobacco, alcohol, and illicit substance use, as well as sexual history data, were entered. This patient-generated data for SH was compared to that already present within the EHR as depicted in Table 5. There were very few instances where patient-generated data did not concur with pre-existing data in the EHR. Only one new data element was found in both the tobacco and alcohol use portions of the SH domain. In the illicit substance portion, three instances of new information were noted. Interestingly, 11 instances of new information were observed for sexual history.

Table 4 – Past Surgical History

Category	n (%)
Duplicate Diagnosis	53 (23.0)
Duplicate Low Granularity Diagnosis	120 (51.9)
New Low Granularity Diagnosis	23 (10.0)
New False Diagnosis	7 (3.0)
New Diagnosis with Utility	26 (11.3)
Duplicate Entry	2 (0.9)

Results of physician ratings were compared between Past Medical and Surgical History diagnoses (Tables 3 & 4) using post-hoc chi-square analysis. Notable differences between medical and surgical diagnoses were observed in New False Diagnoses (18.9 vs. 3.0%) and New Diagnoses with Utility (18.4 vs. 11.3%) with significant results (χ^2 = 63.36, p<0.0001). Results were also compared by gender, age (≥50), and visit type in post-hoc analysis. When combining both medical and surgical diagnoses, we observed a difference in physician ratings between male and female patient-generated diagnoses, with the largest difference observed in the rate of New Low Granularity Diagnoses (13.7 vs. 6.6%, χ^2 = 11.17, p=0.0247). Similarly, the difference in physician ratings of diagnoses generated at New versus Return visit types approached significance (χ^2 = 9.35, p=0.0529), especially with Duplicate Diagnoses and New False Diagnoses (13.6 vs. 26.8% and 23.7 vs 12.4%). Age did not correlate with any difference in physician rating of patient-generated medical or surgical diagnoses (χ^2 = 2.42, p=0.659).

Table 5 – New Social History Data Elements

Category	n (%)
Tobacco	1 (5.9)
Alcohol	1 (5.9)
Illicit Substance	3 (17.6)
Sexual History	11 (64.7)
Total	17

Discussion

This study summarizes the quality and potential value of data generated with web and tablet-based data entry of health history information by patients. Previous studies have found clinical utility in patient-generated data related to chronic disease management, such as entering home blood pressure measurements using an online patient portal [6] and family history [8]. This study provides an analysis of more comprehensive patient-generated data.

Over half of patient-generated medical diagnoses and nearly 75 percent of surgical diagnoses were duplicates of those already recorded within the PMH and PSH portions of the EHR. This is consistent with other studies including Manaktala et al. and Okura et al which found good concordance between patient self-reported data and the medical record [7,10]. A broader study by Tisnado et al. examining a paper-based patient survey compared to the medical record found concordance between both data sources varied considerably. Concordance between diagnoses like "History of acute myocardial infarction" and "Diabetes" were high. Less concordance was evident with more subtle diagnoses such as "Depressed Mood", "High Cholesterol" or history of having had a previous echocardiogram [11]. Our results were similar to these previous studies. Similarly, much of the duplication of previously-recorded diagnoses in this study occurred with cardiovascular health (data not included in results).

Surgical diagnoses mirrored medical diagnoses, in that the vast majority of patient-generated data was a duplication of what already existed in the EHR. Interestingly, the rate of New False Diagnosis was considerably lower for PSH vs PMH (3.0 vs 18.9%). Likely, this is due to clearer knowledge on the part of patients regarding previous procedures. Patient self-knowledge of past surgeries is not a novel finding. In the previously mentioned work by Tisnado, the authors found excellent correlation between the patient survey and medical record for the diagnosis "History of coronary artery bypass or angioplasty" and fair agreement for previous cardiac catheterization [11].

The rate of correlation between patient and medical record is encouraging. More interesting, nearly 20 percent of patient-generated medical diagnoses and 11 percent of surgical ones were found to provide a new, sufficiently granular piece of data that was deemed to be likely accurate based on chart review by a content expert. These diagnoses represent discrete data elements not previously being captured in the PMH or PSH sections of the EHR. However, there was some evidence found within the clinical narratives, laboratory values, or imaging data contained within the institutional EHR. The most frequent patient-generated diagnoses that met criteria related to mental-health, specifically anxiety and depression. Hulse et al. found a similar phenomenon during the introduction of a patient-facing health history tool at their institution [12]. Previous literature comments on the prevalence of mental health diagnoses in patients with gastrointestinal disease [13]. The implications of uncaptured data regarding mental health conditions in this population of patients in a single clinic requires further inspection and analysis.

The failure to capture diagnoses as discrete data elements within the EHR has wide-ranging implications. As the EHR is increasingly utilized for secondary use, the need for complete data is more important. Tate et al. describes data quality as a 'sine qua non' for EHR use in randomized control trials [14]. Accurate and complete data is necessary to ensure high quality phenotyping, cohort discovery and recruitment into randomized controlled trials.

An additional 20 percent of patient-generated medical diagnoses and only three percent of surgical ones were deemed both new and sufficiently granular, but no evidence of disease or prior procedure could be found based on chart review. We were limited by the information provided by our institutional EHR. In this era of fragmented care, patients can have missing clinical data that is contained within other institutions or health systems. In some cases, our institutional EHR included clinical narratives, laboratory values, and the results of imaging studies performed outside. It is reasonable to assume that some portion of these patient-generated diagnoses are true and that supporting healthcare data exists within another institution's EHR .

Nearly ten percent of both medical and surgical diagnoses entered were new additions to the chart, but were of such low granularity that they were not clinically meaningful. Further, nearly half of the duplicated medical diagnoses and two-thirds of the duplicated surgical procedures were rated as less-granular versions of diagnoses or procedures already within the PMH and PSH sections of the EHR.

This study also gives an initial insight into the design of the patient-facing tool. Many of the diagnoses listed in the survey provided sufficient granularity without the need for free text entry by the patient (i.e. "Hypertension", "Diabetes mellitus", "Cholecystectomy"). In order to provide a comprehensive list of medical history conditions and surgical procedures that would not be found cumbersome for most patients, the survey was reliant on free text entry in order to provide sufficient granularity for certain diagnoses and procedures. For instance, when "Fracture or Broken Bone" or "Cancer" was selected, the tool relied on patient free text to specify the anatomic site or type. When "GI surgery" was selected, the tool relied on free text entry to specify the site and procedure. An overwhelming majority of patient-generated diagnoses did not include any free text entries associated with them. This was a major limitation of our study. It can be reasonably inferred that patients will not provide further granularity to diagnoses beyond that which is suggested to them on the Health History tool. Further work needs to be done to elucidate reasons why most patients are not engaging with the tool with free text entry. It may be that during the clinic intake process, patient-generated diagnoses from the tool need immediate addressing and clarification by clinic staff. This would require a major shift in intake workflow, most of which is currently being done by medical assistants. This also is an opportunity to reflect on the design of the tool, specifically in regards to the granularity of the diagnoses provided. Perhaps a drop-down menu approach would provide greater ease of use and ensure greater granularity compared to free text entry by patients. Another limitation was the low response rate of patients eligible to use the tool (34.2%). Clearly, buy-in from clinic personnel needs to be sought prior to introduction of these patient-facing technologies in order to address inevitable changes in workflow. This is being done in further iterations of this work.

One important limitation of the current study and an important area to further characterize is the amount of time required for providers to review patient-entered data. While there is an assumption that having patients enter their own data brings efficiencies, the time saved with patient-entry must be counter-balanced by the time to review and potentially correct and verify patient data. This question could likely be answered with an additional observational study or possibly a log file analysis of provider use of the EHR and consumption of patient entered data.

A significant portion of the patient-generated Social History data matched that which was already present within the EHR. There were very few instances within tobacco, alcohol, and illicit substance use history where patient-generated data was found to provide new information. Likely, this is related to the fact that well-established structured data fields have been in existence for these topics and have been utilized preferentially for some time over unstructured free text [15]. In the domain of sexual history, eleven instances of new information occurred. In nine of these instances, no previous sexual history had been recorded and the patient-generated data represented the first present within the EHR. Thirty nine charts (78 percent) contained previous provider-entered sexual history data. This mirrors previous work by Goyal et al. who observed 82 percent of patient charts containing a sexual history in their cohort of emergency department patients. We also found similar results in that when sexual history had been previously entered by the provider, it had a high rate of concordance with the patient-generated sexual history data [16].

Conclusion

Patient-generated health data is becoming more prevalent in healthcare. We are starting to see the EHR as a dynamic instrument in which patients not only access health data, but are engaged in data input as well. In this pilot study, we observed and graded patient-generated health data obtained during the rollout of a comprehensive health history tool via two patient-facing modalities. While the majority of patient-entered diagnoses were already captured in the EHR, nearly 20 percent of patient-generated medical diagnoses and eleven percent of surgical ones were found to be new, granular, and supported by other data within the chart. These results should be met with guarded optimism. Successful implementation of patient-facing technologies requires significant front-end work. IT experts, informaticists, and clinicians need to be engaged to design a suitable tool that combines ease-of-use with granular data capture. Provider and non-provider clinic personnel buy-in is also necessary early in the process, as these technologies will highly affect existing clinic workflows. Future work will include engaging both patients and clinicians to further optimize the user interface and workflows. Ultimately, it appears that patient-entered health history data has the potential of improving clinical data in the EHR and improving these approaches could potentially enable easy, accurate, granular data entry valuable to clinicians and researchers.

Acknowledgments

This research was supported by the University of Minnesota Department of Surgery, University of Minnesota Physicians, and Fairview Health Services.

References

[1] Davis Giardina T, Menon S, Parrish DE, Sittig DF, Singh H. Patient access to medical records and healthcare outcomes: a systematic review. J Am Med Inform Assoc 2014 Jul-Aug;21(4):737-741.

[2] Otte-Trojel T, de Bont A, Rundall TG, van de Klundert J. How outcomes are achieved through patient portals: a realist review. J Am Med Inform Assoc 2014 Jul-Aug;21(4):751-757.

[3] Wu RR, Orlando LA, Himmel TL, Buchanan AH, Powell KP, Hauser ER, et al. Patient and primary care provider experience using a family health history collection, risk stratification, and clinical decision support tool: a type 2 hybrid controlled implementation-effectiveness trial. BMC Fam Pract 2013 Aug 6;14:111-2296-14-111.

[4] 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.

[5] Forjuoh SN, Ory MG, Wang S, des Bordes JK, Hong Y. Using the iPod Touch for Patient Health Behavior Assessment and Health Promotion in Primary Care. JMIR Mhealth Uhealth 2014 Mar 21;2(1):e14.

[6] Sun J, McNaughton CD, Zhang P, Perer A, Gkoulalas-Divanis A, Denny JC, et al. Predicting changes in hypertension control using electronic health records from a chronic disease management program. J Am Med Inform Assoc 2014 Mar-Apr;21(2):337-344.

[7] Manaktala S, Rockwood T, Adam TJ. Validation of pre-operative patient self-assessment of cardiac risk for non-cardiac surgery: foundations for decision support. AMIA Annu Symp Proc 2013 Nov 16;2013:931-938.

[8] Wu RR, Himmel TL, Buchanan AH, Powell KP, Hauser ER, Ginsburg GS, et al. Quality of family history collection with use of a patient facing family history assessment tool. BMC Fam Pract 2014 Feb 13;15:31-2296-15-31.

[9] Vawdrey DK, Wilcox LG, Collins SA, Bakken S, Feiner S, Boyer A, et al. A tablet computer application for patients to participate in their hospital care. AMIA Annu Symp Proc 2011;2011:1428-1435.

[10] Okura Y, Urban LH, Mahoney DW, Jacobsen SJ, Rodeheffer RJ. Agreement between self-report questionnaires and medical record data was substantial for diabetes, hypertension, myocardial infarction and stroke but not for heart failure. J Clin Epidemiol 2004 Oct;57(10):1096-1103.

[11] Tisnado DM, Adams JL, Liu H, Damberg CL, Chen WP, Hu FA, et al. What is the concordance between the medical record and patient self-report as data sources for ambulatory care? Med Care 2006 Feb;44(2):132-140.

[12] Hulse NC, Ranade-Kharkar P, Post H, Wood GM, Williams MS, Haug PJ. Development and early usage patterns of a consumer-facing family health history tool. AMIA Annu Symp Proc 2011;2011:578-587.

[13] Shah E, Rezaie A, Riddle M, Pimentel M. Psychological disorders in gastrointestinal disease: epiphenomenon, cause or consequence? Ann Gastroenterol 2014;27(3):224-230.

[14] Tate AR, Beloff N, Al-Radwan B, Wickson J, Puri S, Williams T, et al. Exploiting the potential of large databases of electronic health records for research using rapid search algorithms and an intuitive query interface. J Am Med Inform Assoc 2014 Mar-Apr;21(2):292-298.

[15] Chen E, Garcia-Webb M. An analysis of free-text alcohol use documentation in the electronic health record: early findings and implications. Appl Clin Inform 2014 Apr 16;5(2):402-415.

[16] Goyal M, McCutcheon M, Hayes K, Mollen C. Sexual history documentation in adolescent emergency department patients. Pediatrics 2011 Jul;128(1):86-91.

Address for correspondence

Genevieve B. Melton, M.D., M.A. Core Faculty, Institute for Health Informatics, and Associate Professor, Department of Surgery, University of Minnesota, MMC 450, 420 Delaware Street SE, Minneapolis, MN, 55455, USA.

Email: gmelton@umn.edu.