Bias in Recording of Body Mass Index Data in the Electronic Health Record

Susan Rea, PhD¹, Kent R. Bailey, PhD², Jyotishman Pathak, PhD², Peter J. Haug, MD¹
¹Intermountain Healthcare, Salt Lake City, UT; ²Mayo Clinic, Rochester, MN

Abstract

The relationship between patient disease status and the presence or absence of body mass index (BMI) data in the electronic health record (EHR) has not been characterized. We conducted a descriptive study of the completeness of BMI data for three patient cohorts. Cross-sectional descriptions of BMI presence rates per patient were compared between a cohort having at least one ICD-9-CM code for diabetes mellitus (DM) versus a cohort with no diagnosis constraints. Conversely, frequencies of encounter diagnoses were compared among subgroups having BMI recorded or not in both cohorts described and a third cohort having DM codes from a separate organization’s EHR. The data demonstrate a correlation with presence of BMI and higher disease status. This effect may bias the cohort average BMIs, which appear higher than expected. When EHR BMI data are repurposed for research, biases in the selective recording of BMI may affect the results.

Introduction

The quality of electronic health record (EHR) is critical to its secondary usability for health research and improvement of care.[1] The Strategic Health IT Advanced Research Projects secondary-use consortium (SHARPn) is designing and developing open-source tools, services and software for EHR data normalization and high-throughput phenotyping services.[2, 3] We are also investigating data quality issues and methods to expose and/or mitigate them in the SHARPn architectures. The metrics used to assess the quality of EHR data include accuracy or validity, completeness, and comparability across sources.[4-6] We analyzed the completeness of body mass index (BMI) data in three study cohorts. In this descriptive study, we demonstrate that the completeness of BMI data in the EHR was associated with the disease status of the patient.

Obesity has become a major health problem, prevalent in 35.7% of the U.S. adult population.[7] The U.S. Preventive Services Task Force recommended screening of all adult patients for obesity using BMI measured by the clinician.[8] BMI is also a core measure for documentation in the EHR under the Meaningful Use (MU) stage 1 requirements.[9]

Background

Variation in EHR documentation, in general, can be affected by practice-specific workflow, variations in semantics and granularity, and the perceived clinical relevance of the data. More nuanced factors include: lack of mechanisms to enforce standardized use of the EHR, fragmentation of the data over many EHRs where the patient may be seen, and failure of patients or physicians to document some information.[4, 5, 10]. Completeness of BMI data, in particular, has been shown to vary depending upon the facilities studied.[10-12]

Lin and Haug demonstrated the information value of non-randomly missing diagnostic test results on disease prediction modeling accuracy. For example, chest x-rays are typically not ordered for patients in whom no respiratory problems are clinically suspected.[13] Dean, et al, reviewed 126 outcomes studies using EHR data, including potential problems and statistical approaches to confounding and bias. [14] Discernment of such effects may be less evident in the post-facto secondary use setting where there are limited data and loss of context.[15]
The literature concludes that ambiguities will continue to exist in EHR data, that suitability for intended use – a broad view of data quality – should be rigorously explored, and that remedies include both improved accuracy at the source and statistical adjustment at the point of secondary use. In this study, we describe a systematic bias in the recording of BMI data in the EHRs of two large healthcare delivery organizations.

**Methods**

The three study cohorts described below were subsets of larger data sets defined to study the variation and quality (heterogeneity) of diabetes mellitus (DM) phenotyping data between EHRs at Intermountain Healthcare (IMtn) and Mayo Clinic (Mayo). Both organizations obtained IRB approval as well as data sharing agreements for de-identified data. These denominator criteria were used for all cohorts: records of adults (age 18+) who had at least 2 office visits on separate days during 2009 – 2010, with office visit qualified by CPT codes used for Evaluation and Management (E&M) coding for physician office visit billing. Additional criteria or sampling strategy for the current study were as follows:

1. **Cohort IMtn-DM.** A random sample of 2,975* patient records was drawn from 44,723 identified at IMtn having at least one ICD-9-CM code for DM on any ambulatory visit.

2. **Cohort IMtn-all.** A random sample of 3,114* patient records was drawn from over 1 million at IMtn having no restrictions on ICD-9-CM encounter codes.

3. **Cohort Mayo-DM.** All 11,314 patient records having at least one ICD-9-CM code for DM on any ambulatory visit identified at Mayo.

The data used for this study were age, gender, diagnosis codes recorded for patient encounters, height, weight, BMI and measure dates over the period 2009 - 2010. BMI is a function of height and weight:[16]

\[
BMI = \frac{\text{weight (kg)}}{\text{height}^2 \text{ (meters)}} \quad \text{or} \quad BMI = \frac{\text{weight (pounds)} \times 703}{\text{height}^2 \text{ (in)}}
\]

Height and weight were recorded in both EHRs using both metric and U.S. units and were not consistent as to whether one, two or all BMI elements were recorded or generated at one time. In this study, having BMI recorded means a height and a weight observation, not restricted to the same date, or a BMI observation were recorded during the two year period. The IMtn BMI data were generated from structured EHR entry forms used mainly in the ambulatory setting. The aggregate frequency of at least one instance of each 3-digit ICD-9-CM code per patient was generated for subgroups with and without any recorded BMI data in each cohort. At IMtn, the ICD-9-CM codes were extracted from the ambulatory practice management system, which includes inpatient professional services billing. The Mayo diagnosis data were sourced from both inpatient and outpatient encounters in the Decision Support System.

The analysis proceeded in three stages. (1) The two IMtn study cohorts were created to examine the higher than expected average aggregate BMI observed in preliminary data. Random samples of equivalent size (~3,000) were generated, and the completeness and aggregate average were compared between the IMtn-DM and IMtn-all cohorts. (2) To further explore the higher rate of completion found in the IMtn-DM cohort, the frequencies of ICD-9-CM code categories were compared for subgroups having BMI recorded versus missing, in both cohorts. (3) The higher frequencies of diseases, found in both subgroups having BMI recorded, were then examined in the Mayo-DM cohort.

**Results**

A description of the study cohorts and the average and median per patient median BMI are shown in Table 1. As expected, IMtn-DM and Mayo-DM contain older patients, on average, than IMtn-all. Comparing the two IMtn
cohorts, BMI measurements are recorded more for patients in IMtn-DM (79%) than for IMtn-all (55%) (p< .0001), and the average BMI is higher for IMtn-DM (33 vs. 29.) (p< .0001) The Mayo-DM average BMI was also 33.

Table 1. Description of cohorts

|                  | IMtn-DM | IMtn-all | Mayo-DM |
|------------------|---------|----------|---------|
| Target population| 44723   | > 1 million | 11334   |
| Sample size      | 2975    | 3107     |         |
| Average age      | 60      | 47       | 62      |
| Male percent     | 51      | 42       | 53      |
| Female percent   | 49      | 58       | 47      |
| Num w/ wt&ht or BMI | 2343   | 1723     | 9816    |
| Prct w/ wt&ht or BMI | 79      | 55       | 87      |
| Average BMI median/patient | 33.0 | 29.2 | 33.0 |
| Median BMI median/patient | 31.8 | 27.9 | 32.0 |

Table 2 shows the differences in disease status between the subgroups for all cohorts. The ICD-9-CM categories shown were the most frequent disease categories occurring at both IMtn and Mayo, based on both hospital and ambulatory administrative data, among the target populations having DM (Table 1.) Frequencies are not directly comparable between IMtn-DM and Mayo-DM as the data were sourced differently as described in the Methods section. Disease status was consistently higher among subgroups in each cohort having BMI recorded compared to missing. There were no apparent trends in the age or sex distributions between those having BMI data recorded versus missing.

Table 2. Demographic and diagnosis differences by BMI recorded or missing

|                  | IMtn-DM | IMtn-all | Mayo-DM |
|------------------|---------|----------|---------|
|                  | YES | NO | YES | NO | YES | NO |
| n = 2784 | n = 191 | n = 2426 | n = 681 | n = 9816 | n = 1518 |
| Average age      | 60  | 62 | 48  | 45 | 62  | 63 |
| Percent female   | 49  | 52 | 58  | 57 | 47  | 47 |
| ICD-9-CM CODE DESCRIPTION | FREQUENCY (% of patients having at least one ICD-9-CM code) |
| 250 Diabetes mellitus | 100 | 100 | 11  | 4 | 100 | 100 |
| 272 Disorders of lipid metabolism | 71  | 45  | 29  | 7 | 78  | 47 |
| 401 Essential hypertension | 72  | 49  | 29  | 10 | 73  | 53 |
| 786 Respiratory system symptoms | 31  | 30  | 23  | 16 | 37  | 21 |
| 719 Disorders of joint | 22  | 12  | 20  | 12 | 32  | 21 |
| 414 Chronic ischemic heart disease | 14  | 15  | 5   | <5 | 28  | 17 |
| 729 Disorders of soft tissues | 18  | 12  | 14  | 7 | 25  | 17 |
| 278 Obesity and hyperalimentation | 8   | 6   | <5  | <5 | 27  | 11 |
| 715 Osteoarthrosis and allied disorders | 37  | 23  | 9   | 5 | 22  | 14 |
| 724 Disorders of back | 15  | 9   | 16  | 8 | 19  | 12 |
Discussion

The data show BMI was recorded for more patients when a study cohort was selected for having a diagnosis of DM (79%) than for an otherwise equivalent cohort not selected on any particular diagnosis (55%). A bias toward recording the BMI on higher disease status patients was also demonstrated when comparing the frequencies of patients having particular diagnoses between subgroups having versus not having a BMI recorded. We also show questionably high average BMI values for both cohorts selected for having a diagnosis of DM (33) and a cohort not constrained by any diagnosis code (29). A BMI of 25-29 is considered overweight; 30-34, obese, moderate risk of mortality; 35-39, obese, high risk; and 40 and higher, morbid obesity. The IMtn-all cohort average BMI fell at the borderline for obesity. The median BMI, at 28, contrasted with a U.S. adult population statistic of 64.3% below a BMI of 30 in the same time period.[7] The median BMI in both DM cohorts (32) appeared slightly high compared to a study of 3,637 patients attending a DM clinic in the United Kingdom, where 54% were below a BMI of 30.[17] The median for self-reported DM and BMI in a large U.S. population based study was approximately 30.[18] Subjects may have self-selected somewhat in this study, with a low response rate to join (5-8%), which may have favored the less healthy.

Instead of interpreting the average BMIs in this study to be slightly higher than those previously reported, which may be true, we recommend using appropriate resampling, analytic or statistical adjustment methods to determine the average BMI for the target populations. Suspected bias should be examined when BMI data are extracted from the EHR for secondary uses in which it will be aggregated, in general. Noël et al. proposed that a solid understanding of the data quality issues can enable adjustments in the analyses, so that flawed but valuable EHR data can be used for clinical research.[11]

We described the heterogeneity of the sources of EHR data for this study. In a separate analysis, diagnosis codes from outpatient encounter records from the IMtn hospital administrative system were combined with the inpatient and ambulatory encounter records from the ambulatory practice administrative system, used for the IMtn data reported in Table 2. The frequencies of diagnosis categories of the IMtn-DM target population (Table 1) were very similar to those of Mayo-DM. The higher frequency of ICD-9-CM 414, chronic ischemic heart disease, in the IMtn-all cohort (15%) versus IMtn-DM (14%) is likely an artifact of the cardiology care process. Monitoring visits may occur at hospital clinics with specialized services, such as electrocardiograms. These encounter diagnosis codes were missed in the data for this study. The IMtn and Mayo data sets were specified to the same data elements but were not specified to the potential contributing sources for particular data elements. Several studies have reported that locating the sources of secondary data in the EHR is difficult at best.[5, 10, 19] EHRs may contain data sourced from various applications and practice sites, which may vary in content and quality. Concurrent work involves collection, collation and comparison of metadata, including provenance, on EHR applications sourcing particular data elements used for the SHARPn heterogeneity studies. When evaluating the completeness of data, we also missed the rich source of data in clinical text documents. Chan et al. recommend more use of natural language processing (NLP) to extract narrow and consistently defined data from clinical documentation, which is also a goal within SHARPn.

The recording of BMI data may improve over time with MU and other health care quality initiatives. However, unintended consequences of quality measure reporting could include EHR recording on selected patient populations. Stage 1 MU requires BMI documented on 50% of a provider’s patients. Providers may continue to focus on patients for whom BMI is deemed clinically relevant. Additionally, a CMS core quality measure targeting adult weight screening (BMI) and follow-up (NQF 0421)[1], lists the CPT4 codes that define the denominator. We have examined the patients with DM diagnosis codes that visited Intermountain Healthcare who were not selected for the target population used in this study. Preliminary results suggest that less insurance coverage, lower disease status and encounters with specialists may have precluded patient selection. The NQF 0421 denominator specification was less restrictive than ours, but coding practices and patient presentation at different sites of care may introduce bias in the patients targeted for measures, and secondarily, the recording of required data such as BMI.

This study confirms a need for methods to expose EHR data quality issues that have potential for bias and confounding in a secondary research data set. Statistical adjustments appropriate to the use case may proceed from this information. Further research on the effects of the completeness of EHR data, as well as contributing patient
selection biases and heterogeneity among the various data sources and sites, will help leverage truly useful secondary EHR data.

Acknowledgements

The authors would like to thank Ms. Christina M. Wood-Wentz, Mayo Clinic, for data preparation and metadata. Intermountain data were made possible by the Enterprise Data Warehouse. This manuscript/abstract was made possible by funding from the Strategic Health IT Advanced Research Projects (SHARP) Program (90TR002) administered by the Office of the National Coordinator for Health Information Technology. The contents of the manuscript/abstract are solely the responsibility of the authors.

References

1 Hammond WE, Bailey C, Boucher P, Spohr M, Whitaker P. Connecting information to improve health. Health Aff (Millwood). 2010 Feb;29(2):284-8.
2 Chute CG, Pathak J, Savova GK, et al. The SHARPN project on secondary use of Electronic Medical Record data: progress, plans, and possibilities. AMIA Annu Symp Proc. 2011;2011:248-56.
3 Rea S, Pathak J, Savova G, et al. Building a robust, scalable and standards-driven infrastructure for secondary use of EHR data: The SHARPN project. J Biomed Inform. 2012 Feb 4.
4 Botts T, Hartvigsen G, Chen F, Weng C. Secondary Use of EHR: Data Quality Issues and Informatics Opportunities. AMIA Summits on Translational Science proceedings AMIA Summit on Translational Science. 2010;2010:1-5.
5 Chan KS, Fowles JB, Weiner JP. Review: electronic health records and the reliability and validity of quality measures: a review of the literature. Medical care research and review : MCRR. 2010 Oct;67(5):503-27.
6 Orfanidis L, Bamidis PD, Eaglestone B. Data Quality Issues in Electronic Health Records: An Adaptation Framework for the Greek Health System. Health informatics journal. 2004 March 1, 2004;10(1):23-36.
7 Ogden CL, Carroll MD, Kit BK, Flegal KM. Prevalence of obesity in the United States, 2009-2010. NCHS data brief. 2012 Jan(82):1-8.
8 Screening for Obesity in Adults, Topic Page. 2011 Nov, 2011 [cited Mar 2, 2012]; Available from: http://www.uspreventiveservicestaskforce.org/uspstf/uspsobes.htm
9 Medicare & Medicaid EHR Incentive Program Meaningful Use Stage 1 Requirements Summary. 2010 [cited Mar 2, 2012]; Available from: https://www.cms.gov/EHRIncentivePrograms/Downloads/MU_Stage1_ReqSummary.pdf
10 Parsons A, McCullough C, Wang J, Shih S. Validity of electronic health record-derived quality measurement for performance monitoring. J Am Med Inform Assoc. 2012 Feb 9.
11 Noel PH, Copeland LA, Perrin RA, et al. VHA Corporate Data Warehouse height and weight data: opportunities and challenges for health services research. Journal of rehabilitation research and development. 2010;47(8):739-50.
12 Rose SA, Turchin A, Grant RW, Meigs JB. Documentation of body mass index and control of associated risk factors in a large primary care network. BMC health services research. 2009;9:236.
13 Lin JH, Haug PJ. Exploiting missing clinical data in Bayesian network modeling for predicting medical problems. J Biomed Inform. 2008 Feb;41(1):1-14.
14 Dean BB, Lam J, Natoli JL, Butler Q, Aguilar D, Nordyke RJ. Review: use of electronic medical records for health outcomes research: a literature review. Medical care research and review : MCRR. 2009 Dec;66(6):611-38.
15 Rein A. Finding Value in Volume: An Exploration of Data Access and Quality Challenges; 2012 March, 2012.
16 Keys A, Fidanza F, Karvonnen MJ, Kimura N, Taylor HL. Indices of relative weight and obesity. Journal of chronic diseases. 1972 Jul 1;25(6):329-43.
17 Daousi C, Casson IF, Gill GV, MacFarlane IA, Wilding JP, Pinkney JH. Prevalence of obesity in type 2 diabetes in secondary care: association with cardiovascular risk factors. Postgraduate medical journal. 2006 Apr;82(966):280-4.
18 Bays HE, Chapman RH, Grandy S, Group SI. The relationship of body mass index to diabetes mellitus, hypertension and dyslipidaemia: comparison of data from two national surveys. International journal of clinical practice. 2007 May;61(5):737-47.
19 Kristianson KJ, Ljunggren H, Gustafsson LL. Data extraction from a semi-structured electronic medical record system for outpatients: a model to facilitate the access and use of data for quality control and research. Health informatics journal. 2009 Dec;15(4):305-19.