LETTERS—CONCISE RESEARCH REPORTS

Diabetes Phenotyping Using the Electronic Health Record

J Gen Intern Med 35(12):3716–8
 DOI: 10.1007/s11606-020-06231-0
 © Society of General Internal Medicine 2020

INTRODUCTION

Accurate electronic health record (EHR) identification of diabetes is necessary for promoting best practices, clinical decision support, and population health. However, simple definitions have limitations: laboratory-based definitions may miss patients with well-controlled diabetes, medications would miss diet-controlled patients or erroneously capture those taking a “diabetes medication” for another indication, and International Classification of Diseases (ICD) codes require accurate clinician coding. While prior work has examined laboratory work, diabetes medications, and ICD-9 codes,^{1–3} they have not incorporated newer ICD-10 coding.

This study evaluated whether combination of EHR criteria could create a diabetes patient registry that would be useful to clinical practices. Pitfalls from false positives and negatives were clarified with chart review.

METHODS

We conducted a retrospective cohort study of patients aged 18–85 with inpatient or outpatient encounters from January 1, 2014, to November 15, 2016, to the NYU Langone Health System. We created six clinical phenotype definitions using Epic EHR data: [1] type 2 diabetes ICD-10 code listed at least once in either the problem list, medical history, or encounter; [2] laboratory work (hemoglobin A1C ≥ 6.5); [3] active or discontinued outpatient diabetes medication; [4] presence of any of the above; [5] presence of either [1] or [2]; and [6] presence of either [1] or [3]. The EHR pull identified 62,661 patients who met the ICD-10 code definition, 36,151 patients who met the laboratory work definition, and 30,855 patients who met the diabetes medication definition. Together, these represented 76,114 unique patients. Of these, we sampled 106 charts. We also identified 1000 random charts that met none of the above diabetes definitions; we sampled 99 charts from this latter group. Four physicians manually reviewed the total 200 charts to validate each diabetes definition. Reviews were

completed independently, with an overlap of 51 medical records that demonstrated 88% agreement ($\kappa = 0.43$). Differences were adjudicated by H.M.W. To estimate clinical phenotype definition accuracy using chart sampling, we used the methods described by Rosenman and colleagues⁴ to calculate sensitivity, specificity, and positive predictive value (PPV). Physician review was the gold standard in these calculations. The sensitivity of ICD-10 codes alone was compared to each of the other EHR-based phenotypes using the delta process⁴. Qualitative reviews of false positives and negatives were performed with chart review.

RESULTS

Of the 205 charts selected for review, 46% (94 out of 205) were adjudicated as having diabetes. This included 3 of the 99 charts from the group meeting none of the diabetes definitions. Six charts identified by the laboratory definition EHR pull actually had A1Cs ≥ 6.5 outside the study time period. ICD-10 codes were not more sensitive than laboratory work ($p = 0.17$) but were more sensitive than diabetes medication ($p = 0.03$) (Table 1). Diabetes medication had the lowest sensitivity (24%) and PPV (80%). The fifth algorithm, which identified patients with ICD-10 codes or laboratory work, had a sensitivity of 59% ($p = 0.72$ compared to ICD-10 codes alone) and PPV at 87%.

Qualitative review of false positives revealed that some clinicians coded for diabetes when screening, prior to diagnosis (Table 2). False negative review revealed that while diabetes may be listed on the medical history or problem list, for the algorithm to detect it, it must be associated with an encounter diagnosis. Seven patients with diabetes were not picked up by any algorithms—six were on diabetes medications—but as they were listed as ordered an outside provider, they were not detected by the algorithm.

DISCUSSION

EHR-based diabetes phenotypes have moderate PPV but limited sensitivity, indicating these pulls can miss a significant portion of patients with diabetes. Qualitative review of false negatives revealed that administrative data pulls miss patients whose ICD-10 codes are not associated with an encounter diagnosis or with outside provider ordered diabetes medications.

Prior Presentation Parts of this study were presented in abstract form at the SGIM 2018 Annual Meeting, April 12, 2018, Denver, CO.

Received August 19, 2019

Accepted September 10, 2020

Published online September 18, 2020

Table 1 Sensitivity, Specificity, and Positive Predictive Value of Six Different Diabetes Identification Algorithms

Data element	Number of charts identified by EHR pull	Number verified by MD	Validation with medical record review		
			Sensitivity (95% confidence interval)	Specificity	PPV
Diabetes diagnosis*	82	74	0.53 (0.29–0.76)	0.99 (0.99–1.00)	0.87 (0.78–0.96)
Laboratory work	72	68	0.33 (0.18–0.48)	1.00 (1.00–1.00)	0.94 (0.88–1.00)
Diabetes medication	55	48	0.24 (0.13–0.35)	1.00 (0.99–1.00)	0.80 (0.72–0.88)
Presence of ≥ 1 of the above	106	91	0.60 (0.33–0.87)	0.99 (0.98–0.99)	0.81 (0.73–0.89)
Presence of diabetes diagnosis OR laboratory work	100	85	0.59 (0.33–0.86)	0.99 (0.99–1.00)	0.87 (0.78–0.95)
Presence of diabetes diagnosis OR diabetes medication	93	80	0.55 (0.31–0.79)	0.99 (0.99–0.99)	0.81 (0.73–0.90)

*ICD-10 codes E11.X

EHR, electronic health record; MD, physician; PPV, positive predictive value

Table 2 Qualitative Review of False Positives and False Negatives

Data element	Number of false positives	False positive details	Number of false negatives	False negative details
Presence of either diabetes diagnosis, laboratory work, or diabetes medication	16	• Details outlined below	7	<ul style="list-style-type: none"> • 6: On diabetes medication, but were listed as “historical” medications, indicating they were active medications ordered by a provider outside our institution. - Of these, 4 also had diabetes listed on medical history or problem list, however to be included in the final list with ICD-10 codes, code must be associated with an encounter diagnosis. Of these, 3 also had A1C ≥ 6.5 outside our date specifications • 1: A1C ≥ 6.5 outside our date specifications
Single-criteria algorithms Diabetes diagnosis	8	<ul style="list-style-type: none"> • 1: “NIDDM” diagnosis made during inpatient psychiatry hospitalization, however A1C < 6.5, glucose never > 200 • 4: Diabetes coded when A1C was ordered, but result < 6.5 • 1: Chart only contains encounters for radiology visits. Radiology order states “history of diabetes” but no other information to confirm diagnosis • 2: Possible true positives -1: A1C outside adjudicated time frame ≥ 6.5 -1: notes from outside adjudicated time frame indicate diabetes 	15	<ul style="list-style-type: none"> • 8 detected with other single criteria -Laboratory work: 7 -Diabetes medication: 1 • Remaining 7 undetected by presence of either diabetes diagnosis, laboratory work, or diabetes medication as described above
Laboratory work	1	<ul style="list-style-type: none"> • A1C was obtained by non-PCP prior to a procedure (resulted as 8.4). However, abnormal A1C not acknowledged or addressed. In addition, planned procedure was discontinued, after which there were no more encounters for this patient. Possible that A1C of 8.4 represents lab error. Clinical encounters in chart were only for specialty care 	35	<ul style="list-style-type: none"> • 28 detected with other single criteria -Diabetes diagnosis: 27 -Diabetes Medication: 1 • Remaining 7 undetected by presence of either diabetes diagnosis, laboratory work, or diabetes medication as described above
Diabetes medication	7	<ul style="list-style-type: none"> • 1: “NIDDM” diagnosis made during inpatient psychiatry hospitalization and ordered for insulin, however A1C < 6.5, glucose never > 200. Discharged with metformin • 5 on metformin for other indications -1: pre-diabetes -2: polycystic ovarian syndrome -2: weight gain 1 of these patients also on liraglutide for weight loss • 1: colesevelam for hyperlipidemia 	46	<ul style="list-style-type: none"> • 39 detected with other single criteria -Diabetes diagnosis: 32 -Laboratory work: 7 • Remaining 7 undetected by presence of either diabetes diagnosis, laboratory work, or diabetes medication as described above

While chart review can be time intensive, individual practices could easily replicate our study and make simple modifications to calibrate such a tool. For example, a higher PPV of the ICD-10 algorithm could be achieved requiring a higher number [2 or 3] of ICD-10 codes, though this would decrease sensitivity.

Our study is limited by imperfect reliability among physicians who validated the diabetes diagnoses, small study sample, and single-institution design. Regardless, we found algorithms combining clinical criteria were not more sensitive than ICD-10 codes alone in detecting diabetes in the EHR.

Acknowledgments: Hoang-Long Huynh, MD; Amal Shariff, MD; Namrita Boparai, MD; and Jonven Attia, MD, performed medical record review for adjudication of diabetes.

Himali M. Weerahandi, MD, MPH^{1,2,3,4}
 Leora I. Horwitz, MD, MHS^{1,2,3,4}
 Saul B. Blecker, MD, MHS^{1,2,3,4}

¹Division of General Internal Medicine and Clinical Innovation, Department of Medicine, NYU Grossman School of Medicine, New York, NY, USA

²Department of Medicine, NYU Grossman School of Medicine, 227 E. 30th St, New York, NY 10016, USA

³Division of Healthcare Delivery Science, Department of Population Health, NYU Grossman School of Medicine, New York, NY, USA

⁴Center for Healthcare Innovation and Delivery Science, NYU Langone Health, New York, NY, USA

Corresponding Author: Himali M. Weerahandi, MD, MPH; Department of Medicine, NYU Grossman School of Medicine 227 E. 30th St, New York, NY 10016, USA (e-mail: Himali.Weerahandi@nyulangone.org).

Author Contributions H.M.W., S.B.B., and L.I.H. designed the study. H.M.W. wrote the first draft of the manuscript and performed the statistical analyses. S.B.B. and L.I.H. reviewed and edited the manuscript. H.M.W. is the guarantor of this work and, as such, had full access to all the data in the study and takes responsibility for the integrity of the data and the accuracy of the data analysis.

Funding Dr. Weerahandi was supported by the NYU CTSA grant KL2TR001446 from the National Center for Advancing Translational Sciences, National Institutes of Health, and is currently supported by a grant from the National Heart, Lung, and Blood Institute, National Institutes of Health (K23HL145110).

Compliance with Ethical Standards:

Conflict of Interest: The authors declare that they do not have a conflict of interest.

REFERENCES

1. Spratt SE, Pereira K, Granger BB, et al. Assessing electronic health record phenotypes against gold-standard diagnostic criteria for diabetes mellitus. *Journal of the American Medical Informatics Association : JAMIA*. 2017;24:e121-e128.
2. Richesson RL, Rusincovitch SA, Wixted D, et al. A comparison of phenotype definitions for diabetes mellitus. *Jamia*. 2013;20(e2):e319-e326.
3. Zgibor JC, Orchard TJ, Saul M, et al. Developing and validating a diabetes database in a large health system. *Diabetes Res Clin Pract*. 2007;75(3):313-319.
4. Rosenman M, He J, Martin J, et al. Database queries for hospitalizations for acute congestive heart failure: Flexible methods and validation based on set theory. *J Am Med Inform Assoc*. 2013;21(2):345-352.

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.