



# Natural Language Processing Improves Detection of Nonsevere Hypoglycemia in Medical Records Versus Coding Alone in Patients With Type 2 Diabetes but Does Not Improve Prediction of Severe Hypoglycemia Events: An Analysis Using the Electronic Medical Record in a Large Health System

*Diabetes Care* 2020;43:1937–1940 | <https://doi.org/10.2337/dc19-1791>

Anita D. Misra-Hebert,<sup>1,2</sup> Alex Milinovich,<sup>2</sup> Alex Zajichek,<sup>2</sup> Xinge Ji,<sup>2</sup> Todd D. Hobbs,<sup>3</sup> Wayne Weng,<sup>3</sup> Paul Petraro,<sup>3</sup> Sheldon X. Kong,<sup>3</sup> Michelle MocarSKI,<sup>3</sup> Rahul Ganguly,<sup>3</sup> Janine M. Bauman,<sup>2</sup> Kevin M. Pantalone,<sup>4</sup> Robert S. Zimmerman,<sup>4</sup> and Michael W. Kattan<sup>2</sup>

## OBJECTIVE

To determine if natural language processing (NLP) improves detection of nonsevere hypoglycemia (NSH) in patients with type 2 diabetes and no NSH documentation by diagnosis codes and to measure if NLP detection improves the prediction of future severe hypoglycemia (SH).

## RESEARCH DESIGN AND METHODS

From 2005 to 2017, we identified NSH events by diagnosis codes and NLP. We then built an SH prediction model.

## RESULTS

There were 204,517 patients with type 2 diabetes and no diagnosis codes for NSH. Evidence of NSH was found in 7,035 (3.4%) of patients using NLP. We reviewed 1,200 of the NLP-detected NSH notes and confirmed 93% to have NSH. The SH prediction model (C-statistic 0.806) showed increased risk with NSH (hazard ratio 4.44;  $P < 0.001$ ). However, the model with NLP did not improve SH prediction compared with diagnosis code–only NSH.

## CONCLUSIONS

Detection of NSH improved with NLP in patients with type 2 diabetes without improving SH prediction.

The risk of hypoglycemia (glucose  $<70$  mg/dL) with diabetes treatment is well recognized, and its prevention is considered a “critical component of diabetes management” (1). Episodes of severe hypoglycemia (SH), defined in this study as requiring hospitalization or emergency department visit in patients with type 2 diabetes, can be identified through structured electronic medical record (EMR) or

<sup>1</sup>Department of Internal Medicine and Center for Value-Based Care Research, Cleveland Clinic Community Care, Cleveland Clinic, Cleveland, OH

<sup>2</sup>Department of Quantitative Health Sciences, Lerner Research Institute, Cleveland Clinic, Cleveland, OH

<sup>3</sup>Novo Nordisk, Inc., Plainsboro, NJ

<sup>4</sup>Department of Endocrinology, Endocrinology & Metabolism Institute, Cleveland Clinic, Cleveland, OH

Corresponding author: Anita D. Misra-Hebert, [misraa@ccf.org](mailto:misraa@ccf.org)

Received 6 September 2019 and accepted 12 April 2020

This article contains supplementary material online at <https://doi.org/10.2337/figshare.12116709>.

© 2020 by the American Diabetes Association. Readers may use this article as long as the work is properly cited, the use is educational and not for profit, and the work is not altered. More information is available at <https://www.diabetesjournals.org/content/license>.

administrative data (2,3). However, non-severe hypoglycemia (NSH), not requiring assistance for recovery (1), is often reported by patients during outpatient visits and captured only if an EMR hypoglycemia diagnosis code is entered. Natural language processing (NLP) combines computer science and linguistics methods (4) and can be used to extract nonstructured EMR text data. The improvement of EMR hypoglycemia detection using NLP had been previously reported (5). We aimed to describe capture of NSH in patients with type 2 diabetes using both diagnosis codes and NLP and to create a predictive model for SH events.

## RESEARCH DESIGN AND METHODS

We used a modified version of the algorithm used by Kho et al. (6) to identify patients in the Cleveland Clinic Health System with type 2 diabetes during 2005 to 2017. Hypoglycemia events were identified using ICD-9/ICD-10 codes (Supplementary Appendix 1); events not associated with hospitalization/emergency department visit were categorized as NSH. A randomly selected subset of patient records stratified by years was reviewed to confirm code identification of NSH. The study protocol was approved by the Institutional Review Board at Cleveland Clinic (Cleveland, OH).

Using clinical progress notes, an NLP algorithm was developed. The cTAKES program (7) was used to break down sentences into phrases to identify hypoglycemia-related Unified Medical Language System concepts. Regular expressions were written to classify polarity (event or no event) of phrases by pattern matching. Using word embeddings, the remaining phrases were classified with additional pattern matching (see Supplementary Appendix 2 for detailed descriptions). For each year, 100 events were randomly selected, with 2005 to 2006 combined given the lower number of notes. The phrases were reviewed by practicing clinicians and confirmed through record review of a subset of patients as representing true events. The weighted proportion of actual events in a set of 1,200 notes was used to define the positive predictive value.

We then limited our data set to include patients with at least one primary care or endocrinology department visit

each year and created a Cox proportional hazards prediction model for SH (5 years' duration) using the following variables based upon our previous work (2): sex, race, median income (8), history of comorbidities (cardiovascular disease, congestive heart failure [CHF], depression, other psychiatric disorders, dementia, cognitive impairment, chronic kidney disease [CKD], and alcohol or substance abuse), and time-dependent variables including age, insurance type (Medicare, Medicaid, commercial, or other), glycosylated hemoglobin (HbA<sub>1c</sub>), BMI, and diabetes medications (insulin, sulfonylurea, glucagon-like peptide 1 receptor agonists [GLP-1RAs], sodium–glucose cotransporter 2 inhibitors, dipeptidyl peptidase 4 inhibitors, metformin, and  $\alpha$ -glucosidase inhibitors). We had previously identified NSH by ICD codes as a variable associated with SH but for this study added two variables: ever-history of NSH and history of NSH within the past 3 months detected by codes and NLP. We also included interaction variables for HbA<sub>1c</sub>, with insulin, and with sulfonylureas.

## RESULTS

The number of patients with type 2 diabetes in this data set increased from 12,706 in 2005 to 176,001 in 2017 (210,191 unique patients). A total of 1,177,590 progress notes were processed. Upon clinician chart review, 1,111 of 1,200 randomly selected events classified as NSH by NLP were confirmed (93% positive predictive value). From 2005 to 2017, 10,205 NSH events were captured by codes and 14,763 events by NLP, with overlap of only 5 events. Among 204,517 patients with no codes for NSH, evidence of NSH was found in 7,035 (3.4%) using NLP.

The chart review confirmed NSH episodes were often documented under general diabetes diagnosis categories. The incidence proportion of patients with SH by ICD codes increased from 0.3 to 1.7% and those with NSH increased from 0.4 to 1.3% from 2005 to 2017. When NLP was added, the incidence proportion of patients with NSH increased from 0.8% (2005) to 2.6% (2017).

There were 47,280 patients included to create the prediction model for SH (Supplementary Appendix 3). The models using NSH codes alone and codes plus NLP (Table 1) had C-statistics of 0.812 and 0.806, respectively. The codes plus NLP

model showed increased risk for SH with NSH (ever-history of NSH: hazard ratio [HR] 4.44,  $P < 0.001$ ; and NSH in past 3 months: HR 1.65,  $P < 0.001$ ), black race (HR 1.81;  $P < 0.001$ ), Medicaid insurance (HR 1.35;  $P = 0.008$ ), history of cardiovascular disease (HR 1.58;  $P < 0.001$ ), CHF (HR 2.35;  $P < 0.001$ ), depression (HR 1.28;  $P = 0.03$ ), psychiatric disorders other than depression (HR 1.55;  $P < 0.001$ ), alcohol or substance abuse (HR 1.55;  $P = 0.04$ ), and CKD (HR 1.86;  $P < 0.001$ ). The effect of insulin use on SH was greater at lower HbA<sub>1c</sub> (HR of 3.32 when HbA<sub>1c</sub> is 6% [42 mmol/mol]; HR of 2.04 when HbA<sub>1c</sub> is 9% [75 mmol/mol]). There was decreased risk of SH with higher BMI (HR 0.69 when BMI was 30 kg/m<sup>2</sup>; HR 0.63 when BMI was 35 kg/m<sup>2</sup>) and decreased risk with metformin (HR 0.53;  $P < 0.001$ ) and GLP-1RA (HR 0.36;  $P < 0.001$ ). Sulfonylurea use had a mixed effect on the risk of SH (HR of 1.61 when HbA<sub>1c</sub> was 6% [42 mmol/mol]; HR of 0.69 when HbA<sub>1c</sub> was 9% [75 mmol/mol]). The effect of HbA<sub>1c</sub> varied at the extremes. With a reference HbA<sub>1c</sub> of 6% (42 mmol/mol), the HR for SH was 1.59 at an HbA<sub>1c</sub> of 5% (31 mmol/mol), HR was 0.73 at an HbA<sub>1c</sub> of 7% (53 mmol/mol), and HR was 1.39 when HbA<sub>1c</sub> was 9% (75 mmol/mol). Of note, NSH within the past 3 months was only a significant predictor in the the model including NLP. Receiver operating characteristic curves comparing code-detected NSH, NLP-detected NSH, and code plus NLP-detected NSH models were similar (Supplementary Appendix 4), while Kaplan-Meier curves for probability of being free from SH showed lower probability for code-detected NLP (Supplementary Appendix 5).

## CONCLUSIONS

Building upon previous work identifying SH in our health system, we found that applying NLP to progress notes improved NSH detection and that NSH is a significant predictor for SH. However, SH prediction using NSH diagnosis codes did not improve with adding NLP, possibly related to the strength of the other variables in contributing to the model and possible greater severity of coded NSH as the Kaplan-Meier curves suggest. There was little overlap between NSH events by codes plus NLP, suggesting that while NSH may be reported by patients, providers may not enter a hypoglycemia diagnosis code, highlighting the benefit of NLP.

**Table 1—Prediction models for risk of SH according to method used for capturing NSH events: diagnosis codes only or diagnosis codes plus NLP**

Variable	NSH using diagnosis codes (C index = 0.812)			NSH using diagnosis codes plus NLP (C index = 0.806)		
	Adjusted HR*	95% CI	P value†	Adjusted HR‡	95% CI	P value†
History of NSH in past 3 months	1.116	0.805, 1.546	0.51	1.647	1.239, 2.189	<0.001
Ever-history of NSH	8.872	7.464, 10.546	<0.001	4.441	3.745, 5.268	<0.001
Age	1.004	0.996, 1.012	0.30	1.007	0.999, 1.015	0.08
Sex, male	0.906	0.78, 1.054	0.20	0.866	0.745, 1.007	0.06
Race			<0.001			<0.001
White	Reference			Reference		
Black	1.827	1.527, 2.186		1.813	1.515, 2.169	
Other	1.031	0.741, 1.435		1.060	0.762, 1.475	
Median income (per 1,000 U.S. dollars), based on patient's zip code	0.999	0.995, 1.003	0.57	0.999	0.996, 1.003	0.76
Insurance			0.009			0.008
Medicare	Reference			Reference		
Medicaid	1.359	1.013, 1.822		1.351	1.007, 1.813	
Commercial	0.846	0.676, 1.058		0.846	0.676, 1.058	
Other	0.787	0.582, 1.065		0.772	0.571, 1.045	
HbA <sub>1c</sub> %§			<0.001			<0.001
5	1.483	1.200, 1.832		1.593	1.289, 1.969	
6	Reference			Reference		
7	0.770	0.650, 0.912		0.725	0.611, 0.859	
8	1.007	0.832, 1.219		0.932	0.770, 1.129	
9	1.502	1.207, 1.870		1.385	1.113, 1.724	
BMI, kg/m <sup>2</sup>			<0.001			<0.001
25	Reference			Reference		
30	0.712	0.638, 0.795		0.690	0.618, 0.770	
35	0.644	0.569, 0.729		0.630	0.557, 0.713	
Cardiovascular disease	1.625	1.359, 1.944	<0.001	1.579	1.321, 1.889	<0.001
CHF	2.261	1.821, 2.807	<0.001	2.349	1.893, 2.916	<0.001
Depression	1.264	1.005, 1.589	0.045	1.282	1.02, 1.611	0.03
Other psychiatric disorders	1.486	1.253, 1.763	<0.001	1.549	1.307, 1.835	<0.001
Dementia	1.410	0.905, 2.195	0.13	1.367	0.88, 2.124	0.16
Cognitive impairment	1.000	0.633, 1.58	1.0	1.051	0.667, 1.655	0.83
CKD	1.843	1.529, 2.222	<0.001	1.858	1.54, 2.241	<0.001
Alcohol or substance abuse	1.458	0.96, 2.215	0.08	1.551	1.021, 2.354	0.04
Insulin§			<0.001			<0.001
HbA <sub>1c</sub> 5%	3.108	1.976, 4.888		2.637	1.675, 4.153	
HbA <sub>1c</sub> 6%	3.797	3.063, 4.706		3.323	2.671, 4.134	
HbA <sub>1c</sub> 7%	4.238	3.388, 5.300		3.794	3.024, 4.761	
HbA <sub>1c</sub> 8%	3.297	2.638, 4.121		2.921	2.330, 3.661	
HbA <sub>1c</sub> 9%	2.344	1.846, 2.977		2.037	1.600, 2.594	
Sulfonylurea§			<0.001			<0.001
HbA <sub>1c</sub> 5%	2.934	1.765, 4.878		2.484	1.498, 4.120	
HbA <sub>1c</sub> 6%	1.806	1.434, 2.273		1.610	1.280, 2.026	
HbA <sub>1c</sub> 7%	1.160	0.907, 1.483		1.085	0.849, 1.386	
HbA <sub>1c</sub> 8%	0.885	0.698, 1.123		0.851	0.671, 1.079	
HbA <sub>1c</sub> 9%	0.705	0.533, 0.932		0.694	0.525, 0.917	
GLP-1RA	0.364	0.228, 0.581	<0.001	0.361	0.227, 0.576	<0.001
DPP-4	0.924	0.723, 1.181	0.53	0.869	0.68, 1.112	0.26
SGLT2i	0.539	0.255, 1.142	0.11	0.576	0.272, 1.22	0.15
Metformin	0.551	0.465, 0.653	<0.001	0.532	0.449, 0.63	<0.001
AGI	1.279	0.527, 3.105	0.59	1.373	0.566, 3.335	0.48

AGI,  $\alpha$ -glucosidase inhibitor; DPP-4, dipeptidyl peptidase 4 inhibitor; SGLT2i, sodium–glucose cotransporter 2 inhibitor. \*Adjusted for NSH based on diagnosis code. †Wald test. For HbA<sub>1c</sub>, BMI, insulin, and sulfonylurea, it tests the linearity of the variable, all interactions, and all nonlinear terms. ‡Adjusted for NSH based on diagnosis code and NLP. §The restricted cubic spline term was used. The knots are placed at 6%, 7%, and 8% of HbA<sub>1c</sub>. ||The restricted cubic spline term was used. The knots are placed at 25, 30, and 35 kg/m<sup>2</sup> of BMI.

While it is known that NSH may have significant adverse consequences (9), treatment deintensification is uncommon (10,11). Identifying patients with increased risk for hypoglycemia (12,13) and alerting providers may affect provider behavior to adjust treatment for at-risk patients (13). Our prediction model identifies demographic characteristics and comorbidities predicting increased SH risk. Modifiable predictors include history of NSH and sulfonylurea and insulin use predicting increased risk of SH at lower HbA<sub>1c</sub> levels and metformin and GLP-1RAs predicting lower risk. The HbA<sub>1c</sub> effect is of higher predicted risk overall at very low HbA<sub>1c</sub> levels at which treatment deintensification may be relevant, but also at very high HbA<sub>1c</sub>, at which fluctuating blood sugars may increase SH risk and limit ability to attain a lower HbA<sub>1c</sub>.

While our data set is robust, we recognize these limitations: 1) the EMR from one health system does not capture NSH or SH events outside of the system, and 2) lack of diabetes duration information in EMR for prediction. Future work will focus on notifying providers at the point of care of captured NSH events to consider deintensification, medication adjustments to lower SH risk, or meal and activity timing, especially in patients with multiple comorbidities, to reduce risk of future SH.

**Duality of Interest.** This study was funded by Novo Nordisk, Inc. A.D.M.-H. has received research support from Merck, Novo Nordisk, Inc., Boehringer Ingelheim, and the Agency for Healthcare Research and Quality (K08-HS-024128) within the past 12 months. A.M. reports receiving research funding from Merck, Boehringer Ingelheim, Novartis, and Novo Nordisk, Inc., within the past 12 months. A.Z., X.J., J.M.B., and M.W.K. report

receiving research funding from Merck and Novo Nordisk, Inc., within the past 12 months. T.D.H., W.W., M.M., and R.G. report being employees of Novo Nordisk, Inc., and holding company stock. P.P. and S.X.K. report being employees of Novo Nordisk, Inc., and holding company stock at the time of the study. K.M.P. reports receiving research funding from Novo Nordisk, Inc., and Merck; receiving consulting fees from Novo Nordisk, Inc., Sanofi, Eli Lilly and Company, and Merck; and participating in the speakers bureaus of Novo Nordisk, Inc., Merck, and AstraZeneca within the past 12 months. R.S.Z. reports receiving research funding from Novo Nordisk, Inc., and Merck and participating in the speakers bureaus of Merck and Johnson & Johnson within the past 12 months. No other potential conflicts of interest relevant to this article were reported.

**Author Contributions.** A.D.M.-H., A.M., T.D.H., S.X.K., and K.M.P. researched and analyzed the data, contributed to the study design/conception, and contributed to the drafting and/or critical revision of the manuscript. A.Z. researched and analyzed the data. X.J., W.W., P.P., and R.S.Z. researched and analyzed the data and contributed to the study design/conception. M.M. researched and analyzed the data and contributed to the drafting and/or critical revision of the manuscript. R.G. contributed to the study design/conception. J.M.B. oversaw the project, contributed to the study design/conception, and contributed to the drafting and/or critical revision of the manuscript. M.W.K. researched and analyzed the data and contributed to the drafting and/or critical revision of the manuscript. All authors have reviewed and approved the final version of the manuscript. A.D.M.-H. is the guarantor of this work and, as such, had full access to all of the data in the study and takes responsibility for the integrity of the data and the accuracy of the data analysis.

## References

1. American Diabetes Association. 6. Glycemic targets: *Standards of Medical Care in Diabetes—2020*. *Diabetes Care* 2020;43(Suppl. 1): S66–S76
2. Misra-Hebert AD, Pantalone KM, Ji X, et al. Patient characteristics associated with severe hypoglycemia in a type 2 diabetes cohort in a large, integrated health care system from 2006 to 2015. *Diabetes Care* 2018;41:1164–1171

3. Pathak RD, Schroeder EB, Seaquist ER, et al.; SUPREME-DM Study Group. Severe hypoglycemia requiring medical intervention in a large cohort of adults with diabetes receiving care in U.S. integrated health care delivery systems: 2005–2011. *Diabetes Care* 2016;39:363–370
4. Open Health Natural Language Processing (OHNLP) Consortium. OHNLP Consortium home page. Available from [https://www.ohnlp.org/index.php/Main\\_Page](https://www.ohnlp.org/index.php/Main_Page). Accessed 30 October 2017
5. Nunes AP, Yang J, Radican L, et al. Assessing occurrence of hypoglycemia and its severity from electronic health records of patients with type 2 diabetes mellitus. *Diabetes Res Clin Pract* 2016; 121:192–203
6. Kho AN, Hayes MG, Rasmussen-Torvik L, et al. Use of diverse electronic medical record systems to identify genetic risk for type 2 diabetes within a genome-wide association study. *J Am Med Inform Assoc* 2012;19:212–218
7. Savova GK, Masanz JJ, Ogren PV, et al. Mayo clinical Text Analysis and Knowledge Extraction System (cTAKES): architecture, component evaluation and applications. *J Am Med Inform Assoc* 2010;17:507–513
8. American Community Survey (ACS). American Community Survey home page. Available from <https://www.census.gov/programs-surveys/acs>. Accessed 31 October 2019
9. Brod M, Christensen T, Thomsen TL, Bushnell DM. The impact of non-severe hypoglycemic events on work productivity and diabetes management. *Value Health* 2011;14:665–671
10. McAlister FA, Youngson E, Eurich DT. Treatment deintensification is uncommon in adults with type 2 diabetes mellitus: a retrospective cohort study. *Circ Cardiovasc Qual Outcomes* 2017;10:e003514
11. Maciejewski ML, Mi X, Sussman J, et al. Overtreatment and deintensification of diabetic therapy among medicare beneficiaries. *J Gen Intern Med* 2018;33:34–41
12. Karter AJ, Warton EM, Lipska KJ, et al. Development and validation of a tool to identify patients with type 2 diabetes at high risk of hypoglycemia-related emergency department or hospital use. *JAMA Intern Med* 2017;177:1461–1470
13. Vimalananda VG, DeSotto K, Chen T, et al. A quality improvement program to reduce potential overtreatment of diabetes among veterans at high risk of hypoglycemia. *Diabetes Spectr* 2017;30:211–216