**Appendix 1. ICD-9/10 Codes Used to Identify Hypoglycemia**

**ICD-9 codes:**

* 251.0 (Hypoglycemic coma)
* 251.1 (Other specified hypoglycemia)
* 251.2 (Hypoglycemia, unspecified)

**ICD-10 codes:**

* E08.641 (Diabetes mellitus due to underlying condition with hypoglycemia with coma)
* E11.641 (Type 2 diabetes mellitus with hypoglycemia with coma)
* E11.649 (Type 2 diabetes mellitus with hypoglycemia without coma)
* E13.64 (Other specified diabetes mellitus with hypoglycemia)
* E13.641 (Other specified diabetes mellitus with hypoglycemia with coma)
* E13.649 (Other specified diabetes mellitus with hypoglycemia without coma)
* E16.0 (Drug-induced hypoglycemia without coma)
* E16.1 (Other hypoglycemia)
* E15 (Nondiabetic hypoglycemic coma)
* E16.2 (Hypoglycemia, unspecified)

**Appendix 2. Hypoglycemia Natural Language Processing (NLP) Algorithm**

Using the full text from 1,177,590 clinical notes for the study population in the appropriate time range, the cTakes program was used for initial processing. Namely, the first step of the algorithm split each note into sentences/phrases using the default sentence detector.

Next, sentences/phrases were filtered to those containing reference to a hypoglycemia-related Unified Medical Language System (UMLS) concept, which were found by searching all UMLS atoms (ie, over all vocabularies) containing the word “hypoglycemia,” resulting in 319,282 sentences/phrases from 219,898 distinct clinical notes.

Each sentence was then processed through [SUTime](https://nlp.stanford.edu/software/sutime.shtml),1 which is a text annotator built to identify temporal phrases. This step was done to accurately identify not only the occurrence of the event, but when the event took place. For example, if the text said “patient had hypoglycemia last month,” the reference date could then be assigned one month prior to when the note was written. This annotator was also leveraged to identify blood-glucose measurement ranges mentioned as, for example, “…in the 60’s,” which was then noted discretely as a measurement in the range of 60 to <70. Overall, there were 89,710 temporal annotations found based on 66,300 sentences from 56,947 clinical notes.

Next, the polarity of sentences was iteratively classified by pattern matching with regular expressions into “no event,” “severe event,” or “mild/moderate (NSH) event.” First, a sentence was classified as “no event” if it satisfied the associated regular expression. If it was not classified as “no event,” then the detection of the “severe event” regular expression was assessed. Finally, if it was neither “no event” nor “severe event,” then the sentence was classified as “mild/moderate” (NSH) if: the associated regular expression was detected, the SUTime reference date was prior to the note date, or the blood-glucose range was less than 70. As a result, there were 159,329 sentences classified as “no event” from 124,848 notes, 9037 sentences classified as “severe event” from 7811 notes, and 68,974 sentences classified as “mild/moderate” from 56,866 notes.

To classify the polarity of the remaining 83,845 sentences (from 67,320 notes), a [Word2vec](http://docs.h2o.ai/h2o/latest-stable/h2o-docs/data-science/word2vec.html)2 model was developed. This framework maps raw text into vector (coordinate) space based on the hidden layer of a neural network model, whose design is set up by indicating the occurrence of pairs of words within a specified window (in this case, the window was 5 words). This window is referred to as the *context window,* which is what allows the meaning/context of words to be maintained. Therefore, the resulting coordinates of the words are such that words that mean similar things are likely to be closer together in Euclidean distance. As pre-processing steps, all numbers and punctuation were removed from the text strings, all letters were converted to lower case, and each word was *lemmatized* to convert words of equivalent meaning to equivalent text strings. A word vector with 100 components (dimensions) was obtained from the model for each word in the corpus. Each component was then averaged over all words within a given sentence/phrases to obtain sentence vectors. Using the class labels defined in the pattern matching (no event, severe event, mild/moderate event), grids of hyperparameters for random forest, gradient boosting machine (GBM), and neural network models were tuned by 10-fold cross validation. The model producing the lowest error rate was chosen as the final model. In this case, the best model was a GBM which has a cross-validated mean per-class error rate of 0.093 (0.02 no event, 0.2 severe, 0.06 mild/moderate). In all models, outcome classes were balanced during model training. The sentence vectors for the unclassified sentences were then evaluated in the model and the class with highest probability became the label. Of the final set of labeled sentences/phrases, only those classified as a mild/moderate event, obtained from an outpatient note, and not within 7 days of a discrete event were kept.

Finally, to tease out duplicate and/or repeat events, the earliest date of an exact sentence/phrase was considered for a given patient. Additionally, events on the same calendar date with differing sentences were combined into a single event. The final set of events were then fine-tuned, in an iterative process, by manual examination of common patterns by study personnel. In this step, group members met to identify remaining obvious patterns that could be excluded from the set of events, in which subsequent regular expressions were written to do so.

**Event statistics for the 1200 NLP-detected NSH notes reviewed:**

Of the 1111 events confirmed, 1022 (92.0%) had the word “hypoglycemia” or “hypoglycemic” or “hypoglycaemia” (916 distinct phrases), 164 (14.8%) had “low blood sugar” or “low blood glucose” or “low sugar” or “low glucose” (141 distinct phrases), and 1110 (99.9%) had at least one of those (996 distinct phrases). Finally, 67 (5.58%) contained a blood-glucose range indictor (e.g. “BG in the 60’s”).

**Common Expression Examples for Natural Language Processing Hypoglycemia Classification Regular expressions**

* NO EVENT

(?i)((hypoglycemia\s{0,}(:|-){0,1}\s{0,}no(s{0,1})($|\s|,|:|.))|(^\s{0,}no(\s|$))|(denies|educa|no\shypo|instruction|counsel)|(pertinent.{0,}neg)|(frequency.{0,}none)|(\?(\s{0,})(:{0,})(\s{0,})(no|none)(\s|$))|(no\s.{1,}hypoglycemia)|(hypoglycemi(c|a)\s(u{0,1}n{0,1})awarenes)|((^|\?|\s)hypoglycemi(a|c)\s(reactions|frequency)(\s|:|-){0,}no)|((^|\s)(no|not|none)\s.{0,}(hypogly|(low\s(blood|bg|glucose|sugar))))|(neg.{0,}\shypo)|(if|watch out|advised|instructed)\s.{0,}(hypogly|(low\s(blood|bg|glucose|sugar)))|(Hypoglycemia:\s{0,}\{YES/NO:48506\}))

* SEVERE EVENT (excluded)

 (?i)(severe|admission|admitted|hospital|emergency|(\s|^)(ED|ER)(\s|$))

* NON-SEVERE EVENT

(?i)(((hypogly|(low\s(blood|bg|glucose|sugar))).{1,}yes)|(hypoglycemia[^A-Z]{0,}$)|((mild|moderate)\s.{0,}(hypogly|(low\s(blood|bg|glucose|sugar))))|(((hypoglycemi(c|a)|(low\s(blood|bg|glucose|sugar)))(\s|:|-){0,}yes)))

1. Chang AX, Manning C. SUTime: A library for recognizing and normalizing time expressions. In: *Proceedings of the Eighth International Conference on Language Resources and Evaluation (LREC’12)*. Istanbul, Turkey: European Language Resources Association (ELRA); 2012:3735–3740. http://www.lrec-conf.org/proceedings/lrec2012/pdf/284\_Paper.pdf. Accessed March 11, 2020.

2. Word2vec — H2O 3.28.1.1 documentation. http://docs.h2o.ai/h2o/latest-stable/h2o-docs/data-science/word2vec.html. Accessed March 11, 2020.

**Appendix 3. Patient Characteristics by Non-Severe Hypoglycemia Detection**

|  | **No NSH events**n (%) or Median [25/75 Interquartile Range] | **ICD-detected NSH**  n (%) or Median [25/75 Interquartile Range] | **NLP detected NSH** n (%) or Median [25/75 Interquartile Range]  | p-value |
| --- | --- | --- | --- | --- |
| n | 46302 | 529 | 449 |  |
| Age | 61.48 [52.68, 69.75] | 61.43 [50.24, 71.55] | 60.30 [51.67, 68.41] | 0.236 |
| Gender = Male | 23669 (51.1) | 186 (35.2) | 225 (50.1) | <0.001 |
| Race |  |  |  | 0.013 |
| White | 34004 (73.4) | 417 (78.8) | 329 (73.3) |  |
| Black | 7784 (16.8) | 80 (15.1) | 84 (18.7) |  |
| Other | 4514 (9.7) | 32 (6.0) | 36 (8.0) |  |
| HbA1c % | 6.60 [6.20, 7.30] | 6.40 [5.90, 7.00] | 7.10 [6.30, 8.40] | <0.001 |
| BMI | 32.22 [28.28, 37.13] | 31.98 [27.94, 37.26] | 31.85 [27.67, 36.92] | 0.338 |
| Insurance |  |  |  | <0.001 |
| Medicare | 19171 (41.4) | 257 (48.6) | 194 (43.2) |  |
| Medicaid | 2723 (5.9) | 44 (8.3) | 34 (7.6) |  |
| Private health insurance | 15293 (33.0) | 161 (30.4) | 157 (35.0) |  |
| Other | 9115 (19.7) | 67 (12.7) | 64 (14.3) |  |
| Median income (in thousand dollars) based on ZIP | 52.23 [42.60, 67.87] | 51.81 [42.70, 66.61] | 50.72 [42.43, 65.04] | 0.068 |
| Charlson Index Score |  |  |  | <0.001 |
| 0 | 14891 (32.2) | 119 (22.5) | 35 (7.8) |  |
| 1 | 15039 (32.5) | 151 (28.5) | 180 (40.1) |  |
| 2 | 7336 (15.8) | 118 (22.3) | 102 (22.7) |  |
| 3 | 4357 (9.4) | 63 (11.9) | 45 (10.0) |  |
| 4 or more | 4679 (10.1) | 78 (14.7) | 87 (19.4) |  |
| Cardiovascular disease | 13372 (28.9) | 230 (43.5) | 144 (32.1) | <0.001 |
| Congestive heart failure | 2195 (4.7) | 37 (7.0) | 31 (6.9) | 0.006 |
| Depression | 6519 (14.1) | 133 (25.1) | 62 (13.8) | <0.001 |
| Other psychiatric disorders | 15407 (33.3) | 289 (54.6) | 142 (31.6) | <0.001 |
| Dementia | 477 (1.0) | 6 (1.1) | 3 (0.7) | 0.729 |
| Cognitive impairment | 1314 (2.8) | 28 (5.3) | 14 (3.1) | 0.003 |
| Chronic kidney disease | 2460 (5.3) | 40 (7.6) | 41 (9.1) | <0.001 |
| Alcohol or substance abuse | 1218 (2.6) | 20 (3.8) | 12 (2.7) | 0.261 |
| Insulin | 8050 (17.4) | 138 (26.1) | 224 (49.9) | <0.001 |
| GLP-1RA | 1781 (3.8) | 34 (6.4) | 43 (9.6) | <0.001 |
| DPP4 | 4437 (9.6) | 49 (9.3) | 72 (16.0) | <0.001 |
| SGLT-2i | 791 (1.7) | 9 (1.7) | 12 (2.7) | 0.294 |
| Metformin  | 28851 (62.3) | 320 (60.5) | 306 (68.2) | 0.027 |
| Sulfonylurea | 10098 (21.8) | 109 (20.6) | 191 (42.5) | <0.001 |
| AGI  | 107 (0.2) | 12 (2.3) | 2 (0.4) | <0.001 |

AGI, alpha-glucosidase inhibitor; BMI, body mass index; GLP-1RA, glucagon-like-peptide-1 receptor agonist; ICD, International Classification of Diseases; NLP, natural language processing; NSH, non-severe hypoglycemia; SGLT-2i, sodium-glucose cotransporter-2 inhibitor

**Appendix 4.** Receiver Operating Characteristic Curves comparing 3 models with the following non-severe hypoglycemia (NSH) variables: Code-detected NSH, natural language processing (NLP)-detected NSH, and code-detected + NLP-detected NSH.



NLP, natural language processing; NSH, non-severe hypoglycemia

**Appendix 5:** Kaplan-Meier curves for probability of being free from severe hypoglycemia



NLP, natural language processing; NSH, non-severe hypoglycemia