

# Automated Diabetes Case Identification Using Electronic Health Record Data at a Tertiary Care Facility

Sudhi G. Upadhyaya, MS; Dennis H. Murphree Jr, PhD; Che G. Ngufor, PhD;  
Alison M. Knight, MS; Daniel J. Cronk, MS; Robert R. Cima, MD;  
Timothy B. Curry, MD, PhD; Jyotishman Pathak, PhD; Rickey E. Carter, PhD;  
and Daryl J. Kor, MD, MSc

## Abstract

**Objective:** To develop and validate a phenotyping algorithm for the identification of patients with type 1 and type 2 diabetes mellitus (DM) preoperatively using routinely available clinical data from electronic health records.

**Patients and Methods:** We used first-order logic rules (if-then-else rules) to imply the presence or absence of DM types 1 and 2. The “if” clause of each rule is a conjunction of logical **and**, **or** predicates that provides evidence toward or against the presence of DM. The rule includes *International Classification of Diseases, Ninth Revision, Clinical Modification* diagnostic codes, outpatient prescription information, laboratory values, and positive annotation of DM in patients’ clinical notes. This study was conducted from March 2, 2015, through February 10, 2016. The performance of our rule-based approach and similar approaches proposed by other institutions was evaluated with a reference standard created by an expert reviewer and implemented for routine clinical care at an academic medical center.

**Results:** A total of 4208 surgical patients (mean age, 52 years; males, 48%) were analyzed to develop the phenotyping algorithm. Expert review identified 685 patients (16.28% of the full cohort) as having DM. Our proposed method identified 684 patients (16.25%) as having DM. The algorithm performed well—99.70% sensitivity, 99.97% specificity—and compared favorably with previous approaches.

**Conclusion:** Among patients undergoing surgery, determination of DM can be made with high accuracy using simple, computationally efficient rules. Knowledge of patients’ DM status before surgery may alter physicians’ care plan and reduce postsurgical complications. Nevertheless, future efforts are necessary to determine the effect of first-order logic rules on clinical processes and patient outcomes.

© 2017 Mayo Foundation for Medical Education and Research. Published by Elsevier Inc. This is an open access article under the CC BY-NC-ND license (<http://creativecommons.org/licenses/by-nc-nd/4.0/>) ■ *Mayo Clin Proc Inn Qual Out* 2017;1(1):100-110



From the Department of Health Sciences Research (S.G.U., D.H.M., C.G.N., J.P., and R.E.C.), Department of Anesthesiology and Perioperative Medicine (A.M.K., T.B.C., D.J.K.), Department of Physiology and Biomedical Engineering (T.B.C.), Department of Information Technology (D.J.C.), Division of Colon and Rectal Surgery (R.R.C.),

Affiliations continued at the end of this article.

**D**iabetes mellitus (DM) is a metabolic disease resulting from abnormal insulin secretion or insulin resistance, or both. Whether DM is type 1 (type 1 diabetes mellitus [T1DM]) or 2 (type 2 diabetes mellitus [T2DM]), it is the leading cause of microvascular complications, myocardial infarction, stroke, congestive heart failure, and end-stage renal disease that often results in premature death or disability.<sup>1</sup> Patients with DM also undergo surgical procedures at a higher rate than do patients without DM.<sup>2</sup>

Metabolic stresses experienced in the perioperative encounter can alter glycemic homeostasis of patients with DM, which then may result in higher rates of perioperative hyperglycemia, postoperative sepsis, impaired wound healing, and ischemia.<sup>3</sup> In addition, previous studies have shown that postoperative complications such as stroke, urinary tract infection, postoperative hemorrhage, transfusions, wound infection, and even death are more common among patients with uncontrolled DM.<sup>4,5</sup> Thus, it is imperative to identify patients who have DM before

their care episode and initiate appropriate care protocols to optimize their glycemic levels.<sup>3,6,7</sup>

In light of the importance of identifying patients with DM before they receive health care, our objective was to develop a highly sensitive and specific automated electronic phenotyping algorithm for such identification. The algorithm would use available data in the electronic health record (EHR) to drive coordinated clinical processes of care. For instance, results from this algorithm would be used by physicians to facilitate 1 or more of the following steps: (1) schedule a regular glucose check, (2) initiate a protocol-driven DM care pathway, (3) initiate a consult with the diabetes service, or (4) ensure insulin availability in the preoperative care environment. Thus, in addition to high classification performance and computational efficiency, the interpretability of the classification process by care providers was prioritized during the development of this DM classifier.

## METHODS

In 2012, planning was initiated for the development of an automated DM identification algorithm at Mayo Clinic, a tertiary care academic medical center with 110 operating rooms and more than 1200 beds in Rochester, Minnesota. Mayo Clinic's Integrated EHR complies with all the definitions of a comprehensive EHR and at its core provides physicians with the ability to capture patients' demographic characteristics and medical history, which includes documenting patients' medical issues in a free-text format in patients' notes.<sup>8</sup> The institution performs approximately 4000 surgical procedures a month, with an estimated 15% to 17% of surgical patients having DM. The surgical practice of Mayo Clinic began a project to standardize the care of patients with DM across the surgical episode. During this process, the practice determined that it would be more efficient, effective, and reliable if the DM diagnosis was captured at a central location in the EHR before the patient arrived on the day of surgery. Clinicians agreed that the data elements needed to determine whether a patient has DM have frequently existed in the EHR, but the lack of clear identification of a DM diagnosis in the EHR was a barrier to implementation.

## Data Sources

Previous studies have suggested that the most accessible and reliable sources for DM cohort identification were (1) *International Classification of Diseases, Ninth Revision, Clinical Modification (ICD-9-CM)* diagnoses codes (Supplemental Appendix 1, available online at <http://www.mcpiqjournal.org>); (2) laboratory test values; and (3) patient medication data, in varying combinations and thresholds.<sup>9,10</sup>

In addition to these 3 primary domains, several studies have indicated that natural language processing (NLP) can considerably increase accuracy and precision during identification of health issues documented in clinical notes.<sup>11,12</sup> Therefore, in the patient notes section of Mayo Clinic's Integrated EHR, a keyword-based search technique (see Supplemental Table, available online at <http://www.mcpiqjournal.org>) was used to detect all cases where a provider makes a positive annotation of DM in a patient's clinical notes, to further improve the accuracy and precision of our algorithm. A typical patient notes section may contain descriptions such as the following:

- The patient denies any chronic medical conditions; however, review of outside health records reveals that at times in the past 5 years, he was medicated for T1DM
- Hyperparathyroidism, C-section, T2DM
- T2DM, last HbA<sub>1c</sub> unknown IFG/early DM2, 131, January 17, 2014
- Pretransplant history dictated into past medical/surgical history:
  1. Renal diagnosis: ADPKD
  2. Recurrence risk? No
  3. Previous transplants? No
  4. Dialysis pretransplant? No
- T2DM
- History of diabetes, history of hypertension in the past, otherwise normal
- T2DM, diet-controlled
- hx gest DM

Metformin, a common medication used not only in DM management but also for other indications (eg, polycystic ovarian syndrome), was included only in the presence of an abnormal laboratory value measurement. The abnormal values were fasting plasma glucose of greater than or equal to 126 mg/dL, random

plasma glucose of greater than or equal to 200 mg/dL, or the presence of a glycosylated hemoglobin test result (hemoglobin A<sub>1c</sub> [HbA<sub>1c</sub>]).

As patients treated with metformin for indications of DM may have HbA<sub>1c</sub> values that are within normal limits, a decision was made by our endocrinology colleagues to classify patients with any value of HbA<sub>1c</sub> and a prescription for metformin in their EHR as having DM.

### Algorithm Design

We used first-order logic (if-then-else) rules to imply the presence or absence of T1DM and T2DM. This rule-based classification model had a series of logical statements using logical **and** operators and logical **or** operators. Beyond modeling actual human decision making with logical statements, the approach based on the if-then-else rule strikes a balance between accuracy and interpretability for general classification problems.<sup>13</sup> An overview of the classification model is provided in [Figure 1](#). With this method, our rule classified a patient as having DM when the patient's EHR contained the following:

1. One or more outpatient diabetes-related ICD-9-CM diagnosis codes  
**or**
2. At least 1 hypoglycemic medication reported during outpatient medication reconciliation  
**or**
3. A combination of metformin use and a laboratory value exceeding the maximum threshold value  
**or**
4. Any positive annotations of DM in the patient's clinical notes

### Existing Phenotype Algorithms

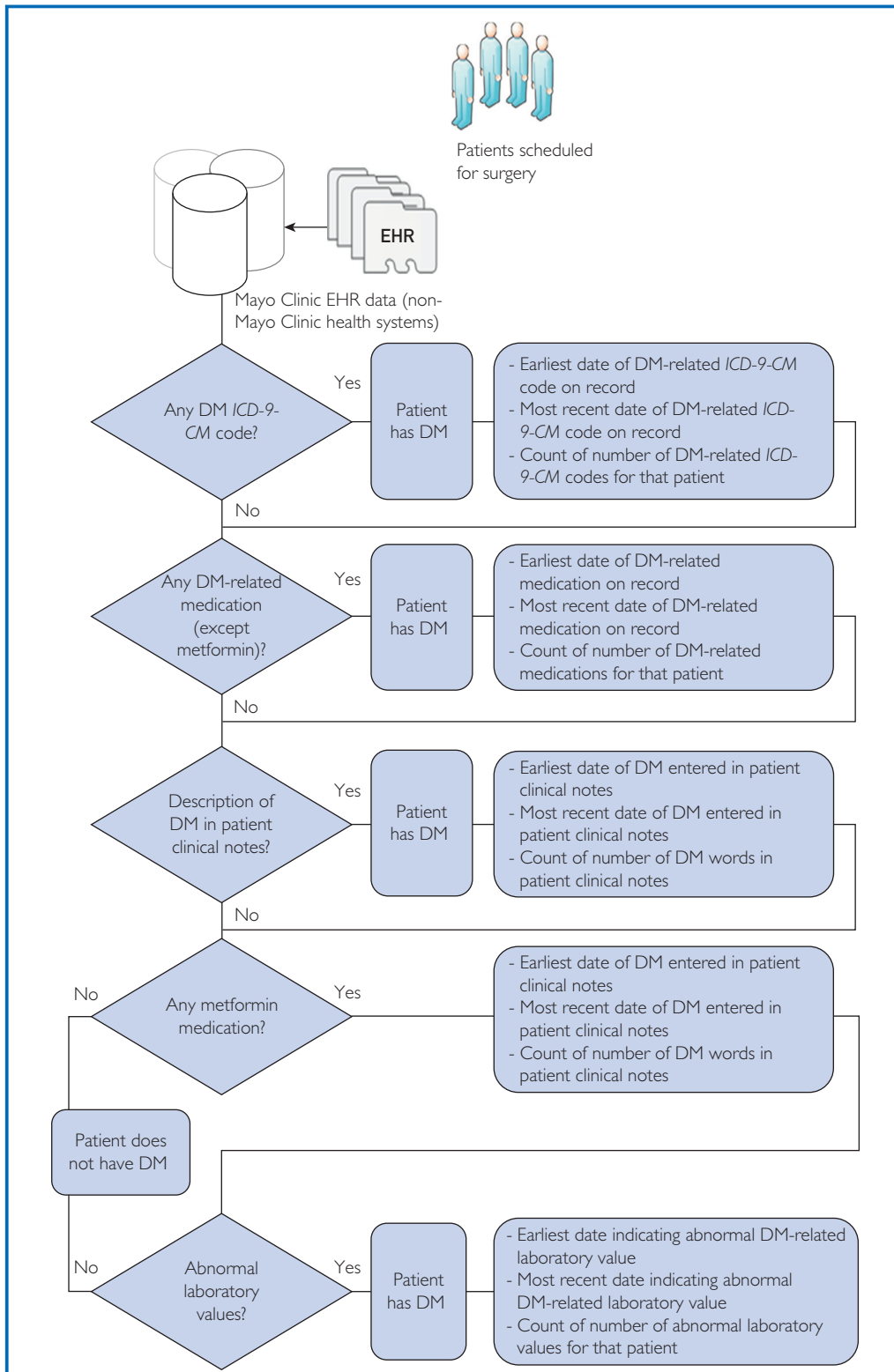
Similar rule-based phenotyping algorithms for the identification of patients with DM have been developed previously. Examples include Electronic Medical Records and Genomics Network (eMERGE), Surveillance Prevention and Management of Diabetes Mellitus (SUPREME-DM), Chronic Condition Data Warehouse (CCW), Durham Diabetes Coalition (DDC), Hemoglobin A<sub>1c</sub> of New York City (HbA<sub>1c</sub> of NYC), and Harvard Medical School.<sup>7,14-24</sup> Although many aspects of these

various algorithms appear similar, considerable differences exist in how these case-definition concepts are operationalized and in the distinctive underlying population characteristics, for which these case-definition strategies were developed.<sup>10,25</sup> For example, the phenotyping algorithm developed by eMERGE focused on identifying only patients with T2DM for large-scale genomic research. SUPREME-DM is a rule-based method developed by a consortium of 11 integrated health systems to identify patients with T1DM and T2DM for research purposes, but this project considered only insured patients. The CCW algorithm was designed to identify T1DM and T2DM among older patients on the basis of Medicare beneficiary data. The DDC designed its DM phenotyping algorithm to identify T2DM cases specifically among racial/ethnic minority groups and low-income patients in Durham County, North Carolina, for public health reasons. HbA<sub>1c</sub> of NYC is a case-definition method designed for mandatory patient HbA<sub>1c</sub> reporting to a public health authority for disease surveillance and tracking. The DM phenotyping algorithm designed by members of Harvard Medical School was targeted toward identifying patients with T1DM and T2DM and was incorporated into an EHR-based public health surveillance and reporting system. Hence, these various DM case-definition methods differ widely on the basis of their stated objectives, and uncertainty remains about their generalizability across various study populations, particularly the cohort of interest in the present investigation—heterogeneous surgical patients.

In light of such limitations, as well as our noted desire to use the results from the developed algorithm for DM determination before the surgical encounter, we were motivated to develop a rule set that could identify accurately the presence of DM in the preoperative environment. This work was performed in collaboration with our endocrinology colleagues to identify a patient who has either T1DM or T2DM at 1 or 2 days before the scheduled surgery date.

### Algorithm Validation

Several authors have adopted various approaches to validate results from phenotyping algorithms. For example, Spratt et al<sup>26</sup>



**FIGURE 1.** Overview of simple first-order rules-based DM phenotyping model. DM = diabetes mellitus; EHR = electronic health record; ICD-9-CM = *International Classification of Diseases, Ninth Revision, Clinical Modification*.

TABLE 1. Accuracy of Mayo Clinic Proposed Method Compared With Preexisting Methods Proposed by Other Authors

Measure	Mayo Clinic proposed	CCW	DDC	SUPREME-DM	eMERGE	HbA <sub>1c</sub> of NYC	Harvard Medical School	Mayo Clinic manual
Sensitivity	0.99 (0.90-0.99)	0.67 (0.63-0.70)	0.92 (0.90-0.94)	0.64 (0.60-0.68)	0.55 (0.51-0.58)	0.48 (0.44-0.52)	0.93 (0.91-0.95)	0.84 (0.82-0.87)
Specificity	0.99 (0.99-1.00)	1.00 (0.99-1.00)	0.94 (0.94-0.95)	0.97 (0.96-0.97)	1.00 (0.99-1.00)	0.99 (0.99-0.99)	0.88 (0.87-0.89)	0.98 (0.98-0.99)
Positive predictive value	0.99 (0.99-0.99)	0.94 (0.93-0.94)	0.98 (0.98-0.98)	0.93 (0.92-0.94)	0.91 (0.91-0.92)	0.90 (0.88-0.90)	0.98 (0.98-0.99)	0.97 (0.96-0.97)
Negative predictive value	0.99 (0.99-1.00)	1.00 (0.98-1.00)	0.78 (0.75-0.80)	0.81 (0.77-0.84)	1.00 (0.98-1.00)	0.94 (0.91-0.96)	0.61 (0.58-0.64)	0.92 (0.90-0.94)
Accuracy	0.99 (0.99-0.99)	0.94 (0.93-0.95)	0.94 (0.93-0.95)	0.91 (0.90-0.92)	0.92 (0.91-0.93)	0.90 (0.89-0.93)	0.89 (0.88-0.90)	0.96 (0.95-0.97)
McNemar $\chi^2$ test (sensitivity)	NA	223.00	43.31	239.00	305.00	348.01	39.09	97.15
P value (sensitivity)	NA	<.01	<.01	<.01	<.01	<.01	<.01	<.01
McNemar $\chi^2$ test (specificity)	NA	1.00	176.02	98.04	1.00	16.20	398.01	41.09
P value (specificity)	NA	.32	<.01	<.01	.32	<.01	<.01	<.01
Total patients identified, n	684	460	815	543	377	350	1043	626

CCW = Chronic Condition Data Warehouse; DDC = Durham Diabetes Coalition; eMERGE = Electronic Medical Records and Genomics Network; HbA<sub>1c</sub> of NYC = Hemoglobin A<sub>1c</sub> of New York City; NA = not applicable; SUPREME-DM, Surveillance = Prevention, and Management of Diabetes Mellitus.

adopted a stratified sampling approach based on Begg and Greenes method, whereas Newton et al<sup>27</sup> adopted an iterative approach where information obtained at each step is used to fine-tune and improve the final phenotype algorithm method. In our study, we followed the iterative approach suggested by Newton et al and evaluated the performance of the proposed algorithm for patients scheduled for surgery during August 2014. Data related to these patients were extracted from the existing legacy Mayo Clinic Unified Data Platform that stores structured, unstructured, and other patient care-related data elements from various sources that support research and quality improvements.<sup>28</sup>

The ascertainment of true DM incidence as the reference standard was created by comparing results from the proposed method with those of the existing process, which is a manual review of the EHR for a DM diagnosis by bedside nurses. When the bedside nurse identifies DM, its documentation is noted in the preoperative evaluation patient flow sheet. This nurse preoperative evaluation patient flow sheet document serves as a surgery intake tool, nursing communication tool, and assessment tool. The nurse documents whether the patient has DM or a history of DM in the flow sheet. This is based on patient response, surgical listing information, the patient's current medication regimen, and a clinical notes review. The nurse synthesizes this information and documents in the EHR if the nurse confirms that the patient has DM. To develop this reference standard, we first considered concordant cases with agreement between the results from the proposed method and the current manual process (eg, both suggest the presence of DM or neither suggests DM). Among such concordant cases, a randomized sample of 100 patients each (ie, a total of 200 concordant cases) was screened by an expert reviewer, who has 35 years of clinical nursing experience, including patient classification, and 16 years of experience in chart abstraction. This review was to confirm whether the patients indeed did or did not have DM (N=200). Second, we considered all discordant cases (ie, a total of 231 discordant cases) where the proposed method disagreed with the findings from the current

manual process. All discordant pairs were screened manually by the independent reviewer to determine whether the patients had DM. Thus, a total of 431 cases were reviewed. Final determination of the presence or absence of DM involved manual review of a patient's EHR by a research nurse specifically trained in the extraction of medical conditions of interest from the EHR. The decision to manually review a random sample of concordant cases was determined a priori and was arbitrarily based on the collective perception of the research team.

### Algorithm Comparison

After developing a reference standard and validating the proposed phenotyping algorithm, we compared the performance of this proposed algorithm with the performance of existing comparator algorithms, as well as current manual imputation of a DM diagnosis by bedside nurses (Table 1).

### Statistical Analyses

Measures of diagnostic accuracy (sensitivity, specificity, positive predictive value, and negative predictive value) and 95% CIs were computed for each case-definition method, and the best method that identified patients with DM most accurately was determined on the basis of McNemar test for sensitivities and specificities.<sup>29,30</sup> All these analyses were conducted using open source R version 3.1.2 (R Foundation for Statistical Computing).<sup>31</sup>

### RESULTS

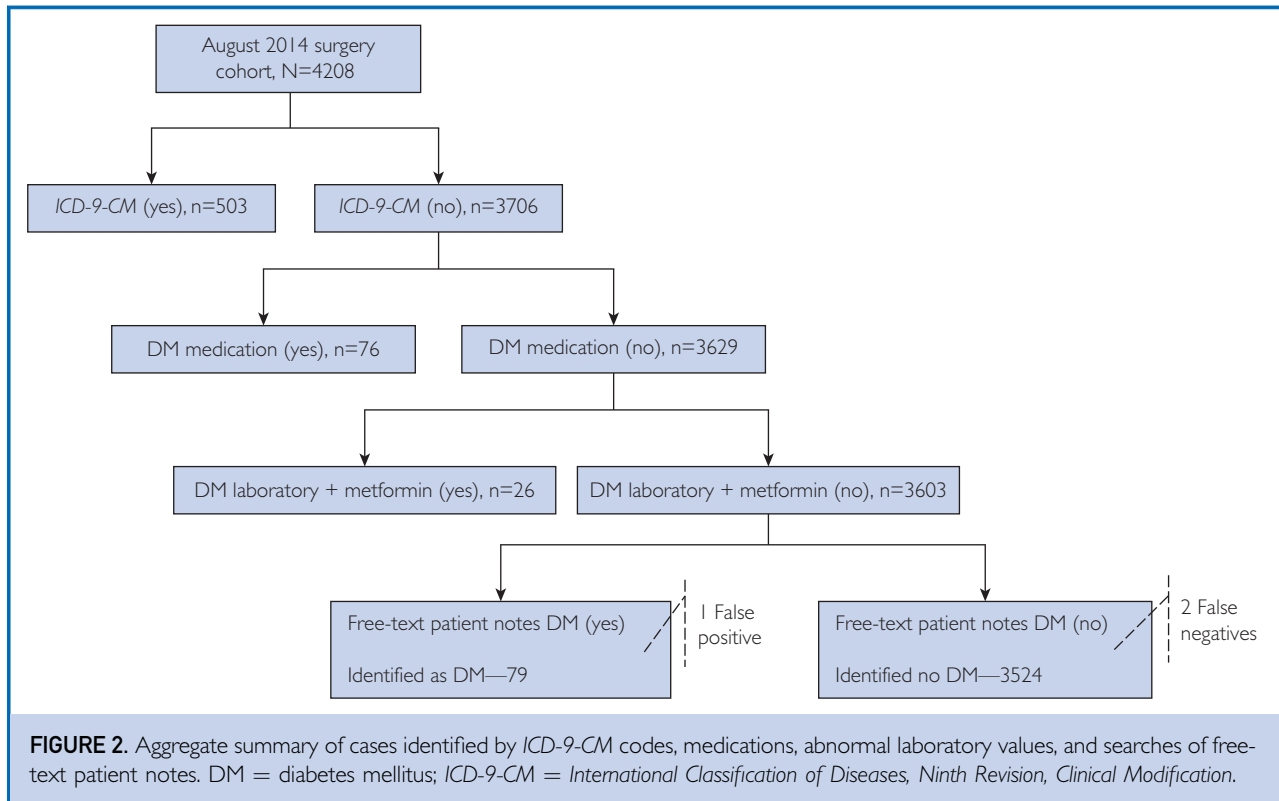
A total of 4208 patients scheduled for surgery in August 2014 were considered in this study. Baseline characteristics of the study population are outlined in Table 2. Our proposed method classified 684 patients as having DM (16.25% of the full study cohort). Of these, 503 (73.53%) were classified as having DM because of the presence of 1 or more outpatient ICD-9-CM codes relating to DM

TABLE 2. Patients' Characteristics at Baseline and Predisposing Factors for DM<sup>a,b</sup>

Characteristic	All patients, (N=4208)	Patients with DM present (n=684)	Patients with DM absent (n=3524)
Age (y)			
<18	501 (11.91)	12 (1.75)	489 (13.88)
18-21	83 (1.97)	1 (0.15)	82 (2.33)
22-29	201 (4.77)	10 (1.46)	191 (5.42)
30-39	337 (8.00)	29 (4.24)	308 (8.74)
40-49	455 (10.81)	94 (13.17)	361 (10.24)
50-65	1256 (29.84)	284 (41.52)	972 (27.58)
66-80	1129 (26.82)	230 (33.63)	899 (25.51)
>80	246 (5.84)	24 (3.51)	222 (6.30)
BMI			
Underweight	378 (8.98)	5 (0.73)	373 (10.58)
Normal	1105 (26.26)	63 (9.21)	1042 (29.57)
Overweight	1199 (28.49)	176 (25.73)	1023 (29.03)
Obese	1489 (35.38)	439 (64.18)	1050 (29.80)
Unavailable	37 (0.88)	1 (0.15)	36 (1.02)
Sex			
Female	2152 (51.14)	298 (43.57)	1854 (52.61)
Male	2056 (48.86)	386 (56.43)	1670 (47.39)
Race			
American Indian/American Native	25 (0.59)	14 (2.14)	11 (0.31)
Non-Hispanic white	3848 (91.44)	615 (90.05)	3233 (91.73)
African American	66 (1.57)	12 (1.68)	54 (1.53)
African	11 (0.26)	1 (0.15)	10 (0.28)
Asian	67 (1.59)	9 (1.23)	58 (1.65)
Native Hawaiian/Pacific Islander	5 (0.12)	1 (0.15)	4 (0.11)
Other	186 (4.42)	32 (4.59)	154 (4.37)

<sup>a</sup>BMI = body mass index; DM = diabetes mellitus.  
<sup>b</sup>Values are n (%).





diagnosis, the number of cases identified by other criteria such as DM-related medications (Supplemental Appendix 2, available online at <http://www.mcpiqjournal.org>), abnormal laboratory values, and metformin use. DM annotations in patient notes are outlined in Figure 2. A single false-positive finding was noted with the keyword searches because of documentation noting “denis [sic] diabetes” instead of “denies diabetes.” Of the other 3524 patients identified as not having DM, 2 patients were incorrectly classified as not having DM by the proposed algorithm (false-negative results). This misclassification was due to a lack of access to source systems that contained DM-related information for these 2 patients. The details are graphically represented using the classification tree diagram of Figure 2.

Of note, the rules implemented in the proposed DM phenotyping algorithm are not mutually exclusive. Nineteen patients were identified as having DM because of the coexistence of metformin and an abnormal DM

laboratory value and also met the criteria for DM on the basis of the presence of a DM-related keyword in their clinical notes. Similarly, the cases identified because of overlapping inclusion criteria are depicted in Figure 3.

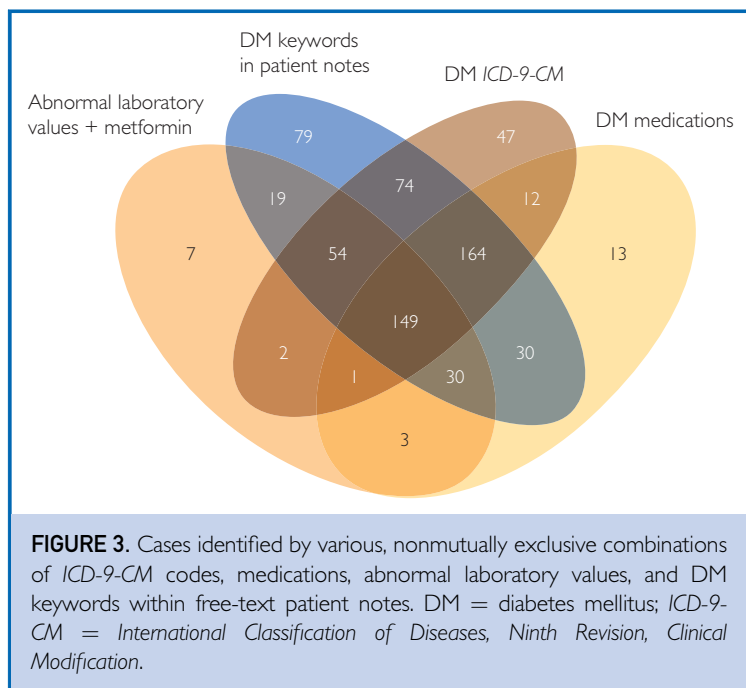
Compared with the 684 patients identified by our proposed approach, the number of DM cases identified by the comparator methods varied considerably: eMERGE, 377 cases; CCW, 460 cases; DDC, 815 cases; SUPREME-DM, 543 cases; HbA<sub>1c</sub> of NYC, 350 cases, and the Harvard Medical School method, 1043 cases. The current manual approach identified 626 patients as having DM (Table 1).

Our proposed method resulted in a sensitivity of 0.9971 (95% CI, 0.9895-0.9996) and a specificity of 0.9997 (95% CI, 0.9984-1.00). The sensitivity of the present phenotyping algorithm exceeded the sensitivity of all other existing methods considered in our study (Table 1). In contrast, the McNemar test for comparison of specificities failed to detect a statistically significant difference between the

specificity on the basis of the present approach and specificities proposed by either CCW (0.99; 95% CI, 0.99-1.00; McNemar test  $\chi^2$ , 1.00;  $P=.32$ ) or eMERGE (1.00; 95% CI, 0.99-1.00; McNemar test  $\chi^2$ , 1.00;  $P=.32$ ). However, statistically significant differences in specificities were observed between our approach and the approaches recommended by DDC, SUPREME-DM, HbA<sub>1c</sub> of NYC, and Harvard Medical School and the manual approach.

**DISCUSSION**

With the rapid adoption of EHRs and the development of innovative data extraction strategies, many institutions have initiated or have completed the development of automated computable case-definition methods. Our results indicate that most of these preexisting methods achieved a high sensitivity of patient identity. Yet, the degradation in the sensitivity of these preexisting methods compared with the current proposed method has many causes. For instance, SUPREME-DM requires that a patient have at least 2 abnormal laboratory values or outpatient ICD-9-CM diagnosis codes not more than 2 years apart (Table 3). The DDC requires 2 abnormal laboratory values not more than 1 year apart. This requirement excludes patients with a single abnormal laboratory value or ICD-9-CM diagnosis codes. Such data result in increased false negatives and reduced



**FIGURE 3.** Cases identified by various, nonmutually exclusive combinations of ICD-9-CM codes, medications, abnormal laboratory values, and DM keywords within free-text patient notes. DM = diabetes mellitus; ICD-9-CM = International Classification of Diseases, Ninth Revision, Clinical Modification.

sensitivity. CCW and HbA<sub>1c</sub> of NYC are single-element case-definition methods. The CCW requires only patient DM-related ICD-9-CM codes, and HbA<sub>1c</sub> of NYC requires only abnormal HbA<sub>1c</sub> values and does not identify patients with other indicators, such as DM-related medications or DM-related ICD-9-CM diagnostic codes and resulting in large false negatives and lower sensitivity.

**TABLE 3. Overview of Differences Among Other Preexisting Methods Proposed by Other Authors**

Existing diabetes mellitus phenotype	Reference	Data elements used				Notes
		Lab	Medications	ICD-9	Patient notes	
CCW	Chronic Condition Data Warehouse <sup>23</sup>			x		Only ICD-9 codes
DDC	Spratt et al <sup>22</sup>	x	x	x		≥2 abnormal lab values, not more than 1 y apart
SUPREME-DM	Desai et al <sup>15</sup>	x	x	x		≥2 abnormal lab values or ICD-9, ≤2 y apart; insured patients only
eMERGE	Kho et al <sup>18</sup>	x	x	x		Mainly designed to identify patients with T2DM and must satisfy ≥2 criteria
HbA <sub>1c</sub> NYC	Chamany et al <sup>14</sup>	x				Only HbA <sub>1c</sub>
Harvard Medical School	Klompas et al <sup>19</sup>	x	x	x		Only abnormal lab value with no metformin requirement

CCW = Chronic Conditions Data Warehouse; DDC = Durham Diabetes Coalition; eMERGE = Electronic Medical Records and Genomics Network; HbA<sub>1c</sub> = hemoglobin A<sub>1c</sub>; HbA<sub>1c</sub> NYC = HbA<sub>1c</sub> of New York City; ICD-9 = International Classification of Diseases, Ninth Edition; lab = laboratory; SUPREME-DC = Surveillance, Prevention and Management of Diabetes Mellitus; T2DM = type 2 diabetes mellitus.



The eMERGE DM algorithm requires at least 2 different DM-related data elements to classify patients as having DM and ignores patients with a single instance of either DM-related ICD-9-CM code or medication **or** metformin and abnormal values. The algorithm does not categorize these patients as having DM, thus lowering its sensitivity. In addition, neither method considered exploring the DM cases mentioned by the physicians in free-text patient notes data. This point further contributed to lower sensitivity among these different comparator methods.

Similar concerns exist when considering the varied specificities of the alternative algorithms. The rules proposed by the SUPREME-DM, DDC, and Harvard Medical School methods resulted in more false positives than with our proposed method because these rules classified patients as having DM solely on the basis of abnormal laboratory values. Concerns about the use of abnormal laboratory values as the singular most important indicator of DM have been raised by several investigators. Sacks et al<sup>32</sup> have suggested that neither random nor fasting glucose concentrations measured in an accredited laboratory should be used as the primary indicator of DM. Instead, they recommended supplementing these criteria with other information. Similarly, Monnier et al<sup>33</sup> stated that no unifying argument supports the accuracy of random or fasting glucose test when identifying patients who may have DM. The methods proposed by SUPREME-DM, DDC, and Harvard Medical School that identified patients with DM through only abnormal laboratory values resulted in large false positives and lower specificity compared with our proposed method.

### Strengths and Limitations

A major strength of this investigation is its inclusion of information from multiple domains relevant to DM. This included ICD-9-CM diagnostic codes, medication information, laboratory findings, and keyword searches mining unstructured text contained within clinical notes.

Several investigators have emphasized the importance of incorporating information buried within a patient's clinical notes when attempting to identify patients with specific

disease conditions.<sup>34-36</sup> Sohn et al<sup>37</sup> noted that rule-based algorithms performed well when NLP-extracted attributes from clinical text were incorporated into the algorithm design. Peissig et al<sup>38</sup> studied cataract cases and asserted that incorporating concepts from clinical notes increased the case identification by a factor of 3 compared with using a sole mode-structured data approach. We used a combination of complex Structured Query Language embedded with Boolean logic ("and" "or") to detect either negation or positive annotation indicative of DM in patient notes (Supplemental Appendix 3, available online at <http://www.mcpiqjournal.org>). Within such a construct, we were able to identify several dominant patterns and develop a set of inclusion and exclusion criteria that indicated the presence or absence of DM. In addition to identifying various affirmation statuses, negation had a notable role in identifying patients who did not have DM and in improving the accuracy of our algorithm. Unlike confirmatory avowals, negations are commonly indicated by such terms as *no*, *NA*, and *unknown*, so initially we started with these 3 basic terms and iteratively refined our algorithm by adding more negation clauses, such as *brothers, one of whom has diabetes type, insignificant for diabetes mellitus, doesn't have a history of diabetes, diabetes screen: fasting glucose = NA*, and *diabetes screen: unknown if ever checked*, and many others during the iterative validation process.

This combination attributed 79 additional cases identified by our algorithm, resulting in performance superior to that of other approaches that did not include unstructured data. Thus, it follows that such a text-parsing NLP algorithm that uses simple keywords to negate or validate the presence of DM will be an EHR system—agnostic and highly portable that can be implemented in other EHR environments.

Inclusion of metformin as an additional requirement for patients with abnormal laboratory values is another distinctive aspect of our algorithm that may have resulted in higher specificity than with other available phenotyping methods. Our method also did not mandate multiple observations per patient at various time intervals. This proposed method has the utility of identifying patients before their surgical encounter rather than after it.

Finally, implementation of the manual review for final determination of DM by a trained study coordinator enhanced our confidence in accuracy regarding the presence or absence of the condition of interest.

Several study limitations merit discussion. First, this investigation represents a single institution's experience in designing, testing, and implementing a DM case-definition algorithm compared with multi-institution case-definition methods, such as eMERGE, SUPREME-DM, and Harvard Medical School. Further validation and confirmation of the present findings in alternate health care settings across various health care institutions are necessary to understand the generalizability of our proposed method. Many health care institutions will not have access to some of the data elements proposed by our case-definition strategy—most notably, the ability to query unstructured data for key terms of interest. Although we were able to validate the results from our algorithm using an independent nurse reviewer, we were unable to completely mask the details to the reviewer during the validation because of the iterative nature of algorithm development.

In addition to these concerns, we acknowledge the potential for bias due to the use of a reference standard and not a criterion standard developed by an independent reviewer to whom results were blinded completely while evaluating the accuracy of our proposed method in our study. However, the CIs around our sensitivity and specificity estimates support the stability of our results to our results and mitigate concerns related to the use of such reference standards during the validation process. Currently, a more formal validation of the performance of our proposed method in a clinical environment at our institution is under way.

Although the use of *International Classification of Diseases, Tenth Edition* codes is the current standard, this study used *ICD-9-CM* codes. This course was taken because the standard at the time of study onset was *ICD-9-CM*. Moreover, the comparator algorithms (DDC, CCW, eMERGE, Harvard Medical College, Supreme-DM) were also developed using *ICD-9-CM* codes. Nonetheless, we acknowledge the use of *ICD-9-CM* codes in our study as a limitation, and we have implemented a

strategy to map the *ICD-9-CM* codes associated with this study to their corresponding *International Classification of Diseases, Tenth Edition* codes.

## CONCLUSION

We developed an efficient and accurate DM phenotypic algorithm that outperformed other available approaches. This algorithm facilitates the early recognition of DM, thereby permitting the implementation of improved workflows for optimal DM-related care in the perioperative period.

## SUPPLEMENTAL ONLINE MATERIAL

Supplemental material can be found online at <http://www.mcpiqjournal.org>. Supplemental material attached to journal articles has not been edited, and the authors take responsibility for the accuracy of all data.

**Abbreviations and Acronyms:** CCW = Chronic Condition Data Warehouse; DDC = Durham Diabetes Coalition; DM = diabetes mellitus; eMERGE = Electronic Medical Records and Genomics; EHR = electronic health record; HbA<sub>1c</sub> = hemoglobin A<sub>1c</sub>; HbA<sub>1c</sub> of NYC = Hemoglobin A<sub>1c</sub> of New York City; *ICD-9-CM* = *International Classification of Diseases, Ninth Revision, Clinical Modification*; MICS = Mayo Integrated Clinical Systems; NLP = natural language processing; SUPREME-DM = Surveillance, Prevention, and Management of Diabetes Mellitus; T1DM = type 1 diabetes mellitus; T2DM = type 2 diabetes mellitus

**Affiliations (Continued from the first page of this article.):** and Robert D. and Patricia E. Kern Center for Science of Health Care Delivery (R.R.C.), Mayo Clinic, Rochester, MN.

**Grant Support:** The work was supported in part by the Mayo Clinic Robert D. and Patricia E. Kern Center for the Science of Health Care Delivery and a grant from the National Institutes of Health (NIH) (grant number R01 GM105688-01). The content is solely the responsibility of the authors and does not necessarily represent the official views of the NIH. The funding sources for this work had no role in the study design; in the collection, analysis, and interpretation of data; in the writing of the report; or in the decision to submit the article for publication.

**Correspondence:** Address to Daryl J. Kor, MD, Department of Anesthesiology and Perioperative Medicine, Mayo Clinic, 200 First St SW, Rochester, MN 55905 ([kor.daryl@mayo.edu](mailto:kor.daryl@mayo.edu)).

## REFERENCES

1. Ponce BA, Menendez ME, Oladeji LO, Soldado F. Diabetes as a risk factor for poorer early postoperative outcomes

- after shoulder arthroplasty. *J Shoulder Elbow Surg.* 2014;23(5):671-678.
2. Ardeshiri M, Faritus Z, Ojaghi-Haghighi Z, Bakhshandeh H, Kargar F, Aghili R. Impact of metabolic syndrome on mortality and morbidity after coronary artery bypass grafting surgery. *Res Cardiovasc Med.* 2014;3(3):e20270.
  3. Dagogo-Jack S, Alberti KGMM. Management of diabetes mellitus in surgical patients. *Diabetes Spectr.* 2002;15(1):44-48.
  4. Krzych LJ, Wybraniec MT. Glycaemic control in cardiac surgery patients: a double-edged sword. *Curr Vasc Pharmacol.* 2015;13(5):578-586.
  5. Marchant MH Jr, Viens NA, Cook C, Vail TP, Bolognesi MP. The impact of glycemic control and diabetes mellitus on perioperative outcomes after total joint arthroplasty. *J Bone Joint Surg Am.* 2009;91(7):1621-1629.
  6. American Diabetes Association. (13) Diabetes care in the hospital, nursing home, and skilled nursing facility. *Diabetes Care.* 2015;38(Suppl):S80-S85.
  7. Handelsman Y, Bloomgarden ZT, Grunberger G, et al. American Association of Clinical Endocrinologists and American College of Endocrinology—clinical practice guidelines for developing a diabetes mellitus comprehensive care plan—2015: executive summary. *Endocr Pract.* 2015;21(4):413-437.
  8. Pickering BW, Herasevich V, Ahmed A, Gajic O. Novel representation of clinical information in the ICU: developing user interfaces which reduce information overload. *Appl Clin Inform.* 2010;1(2):116-131.
  9. American Diabetes Association. Diagnosis and classification of diabetes mellitus. *Diabetes Care.* 2005;28(Suppl 1):S37-S42.
  10. Richesson RL, Rusinovitsh SA, Wixted D, et al. A comparison of phenotype definitions for diabetes mellitus. *J Am Med Inform Assoc.* 2013;20(e2):e319-e326.
  11. Clifford L, Wilson GA, Gajic O, et al. Natural language processing of chest radiograph reports improves the identification of transfusion-related pulmonary complications. *Am J Respir Crit Care Med.* 2013;187:A2218.
  12. Kho AN, Pacheco JA, Peissig PL, et al. Electronic medical records for genetic research: results of the eMERGE consortium. *Sci Transl Med.* 2011;3(79):79re1.
  13. Letham B, Rudin C, McCormick TH, Madigan D. Interpretable classifiers using rules and Bayesian analysis: building a better stroke prediction model. *Ann Appl Stat.* 2015;9(3):1350-1371.
  14. Chamany S, Silver LD, Bassett MT, et al. Tracking diabetes: New York City's A1C registry. *Milbank Q.* 2009;87(3):547-570.
  15. Desai JR, Vazquez-Benitez G, Xu Z, et al; SUPREME-DM Study Group. Who must we target now to minimize future cardiovascular events and total mortality? Lessons from the Surveillance, Prevention and Management of Diabetes Mellitus (SUPREME-DM) cohort study. *Circ Cardiovasc Qual Outcomes.* 2015;8(5):508-516.
  16. Gorina Y, Kramarow EA. Identifying chronic conditions in Medicare claims data: evaluating the Chronic Condition Data Warehouse algorithm. *Health Serv Res.* 2011;46(5):1610-1627.
  17. Hauser JR, Toubia O, Evgeniou T, Befurt R, Dzyabura D. Disjunctions of conjunctions, cognitive simplicity, and consideration sets. *J Mark Res.* 2010;47(3):485-496.
  18. 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(2):212-218.
  19. Klompas M, Eggleston E, McVetta J, Lazarus R, Li L, Platt R. Automated detection and classification of type 1 versus type 2 diabetes using electronic health record data. *Diabetes Care.* 2013;36(4):914-921.
  20. Nichols GA, Desai J, Elston Lafata J, et al; SUPREME-DM Study Group. Construction of a multisite DataLink using electronic health records for the identification, surveillance, prevention, and management of diabetes mellitus: the SUPREME-DM project. *Prev Chronic Dis.* 2012;9:E110.
  21. Pacheco JA, Thompson W, Kho A. Automatically detecting problem list omissions of type 2 diabetes cases using electronic medical records. *AMIA Annu Symp Proc.* 2011;2011:1062-1069.
  22. Spratt SE, Batch BC, Davis LP, et al. Methods and initial findings from the Durham Diabetes Coalition: integrating geospatial health technology and community interventions to reduce death and disability. *J Clin Transl Endocrinol.* 2015;2(1):26-36.
  23. Chronic Conditions Data Warehouse. Condition categories [Internet]. <https://www.cdwdata.org/web/guest/condition-categories>. Accessed October 25, 2016.
  24. Wei WQ, Leibson CL, Ransom JE, et al. Impact of data fragmentation across healthcare centers on the accuracy of a high-throughput clinical phenotyping algorithm for specifying subjects with type 2 diabetes mellitus. *J Am Med Inform Assoc.* 2012;19(2):219-224.
  25. Shivade C, Raghavan P, Fosler-Lussier E, et al. A review of approaches to identifying patient phenotype cohorts using electronic health records. *J Am Med Inform Assoc.* 2014;21(2):221-230.
  26. Spratt SE, Pereira K, Granger BB, et al. Assessing electronic health record phenotypes against gold-standard diagnostic criteria for diabetes mellitus. *J Am Med Inform Assoc.* 2017;24(e1):e121-e128.
  27. Newton KM, Peissig PL, Kho AN, et al. Validation of electronic medical record-based phenotyping algorithms: results and lessons learned from the eMERGE network. *J Am Med Inform Assoc.* 2013;20(e1):e147-e154.
  28. Chute CG, Beck SA, Fisk TB, Mohr DN. The Enterprise Data Trust at Mayo Clinic: a semantically integrated warehouse of biomedical data. *J Am Med Inform Assoc.* 2010;17(2):131-135.
  29. Carter RE. I can't make heads or tails out of what you are saying, so let's just agree to be fair. *Teach Stat.* 2013;35(3):127-130.
  30. McNemar Q. Note on the sampling error of the difference between correlated proportions or percentages. *Psychometrika.* 1947;12(2):153-157.
  31. R Foundation. The R project for statistical computing [Internet]. <https://www.r-project.org/>. Accessed October 25, 2016.
  32. Sacks DB, Bruns DE, Goldstein DE, Maclaren NK, McDonald JM, Parrott M. Guidelines and recommendations for laboratory analysis in the diagnosis and management of diabetes mellitus. *Clin Chem.* 2002;48(3):436-472.
  33. Monnier L, Lapinski H, Colette C. Contributions of fasting and postprandial plasma glucose increments to the overall diurnal hyperglycemia of type 2 diabetic patients: variations with increasing levels of HbA(1c). *Diabetes Care.* 2003;26(3):881-885.
  34. Abhyankar S, Demner-Fushman D, Callaghan FM, McDonald CJ. Combining structured and unstructured data to identify a cohort of ICU patients who received dialysis. *J Am Med Inform Assoc.* 2014;21(5):801-807.
  35. Han D, Wang S, Jiang C, et al. Trends in biomedical informatics: automated topic analysis of JAMIA articles. *J Am Med Inform Assoc.* 2015;22(6):1153-1163.
  36. Scheurweghs E, Luyckx K, Luyten L, Daelemans W, Van den Bulcke T. Data integration of structured and unstructured sources for assigning clinical codes to patient stays. *J Am Med Inform Assoc.* 2016;23(e1):e11-e19.
  37. Sohn S, Kocher JP, Chute CG, Savova GK. Drug side effect extraction from clinical narratives of psychiatry and psychology patients. *J Am Med Inform Assoc.* 2011;18(Suppl 1):i144-i149.
  38. Peissig PL, Rasmussen LV, Berg RL, et al. Importance of multimodal approaches to effectively identify cataract cases from electronic health records. *J Am Med Inform Assoc.* 2012;19(2):225-234.