1 Overview

This document describes the Northwestern University Type 2 diabetes mellitus (T2DM) algorithms for extracting both T2DM cases and T2DM controls from the electronic medical record (EMR). There are two main parts to this document. The first part (Section 2) provides descriptions of the input data elements to be extracted from the EMR, flowcharts, and pseudo-code descriptions of the algorithms. The second part (Section 3) is an installation guide for executable workflows that implement the T2D case and control selection algorithms. These workflows are based on the Konstanz Information Miner (KNIME) data analysis platform.¹

2 Algorithm Descriptions

The case and control selection algorithms require certain patient-level data elements to be extracted from the EMR. This information includes diagnoses, lab results, medication orders, and physician encounter dates. Lists of codes that satisfy various algorithm requirements (including ICD-9 codes, LOINC codes, and RxNorm codes) are provided in tabular form in Appendix A. Additionally, Section 3.1 contains a translation of these data elements into data dictionaries for input into the KNIME workflow implementations.²

¹Questions about the core algorithms should be sent to japacheco@northwestern.edu, while questions about the executable KNIME workflows should be sent to wkt@northwestern.edu.

²See also the T2D study in the elemap online tool. This study contains data elements that were used in a T2D genome-wide association study (GWAS), using a patient cohort derived from the EMR-based algorithm described in this document.
2.1 T2DM Case Selection Algorithm Logic

For the T2D case selection algorithm, the following data elements are required:

1. Counts of T1DM ICD-9 code assignment dates by diagnostic source (Table 3)
2. Counts of T2DM ICD-9 code assignment dates by diagnostic source (Table 4)
3. T1DM medications (i.e., Insulin & Symlin) order or prescription dates – at least the earliest date of Rx (Table 5)
4. T2DM medications order or prescription dates – at least the earliest date of Rx (Table 6)
5. Fasting blood glucose lab values – at least the maximum value (Table 7)
6. Random blood glucose lab values – at least the maximum value (Table 7)
7. HBA1c lab values – at least the maximum value (Table 7)

For the T2D case selection algorithm, the following definitions apply:

- **Abnormal lab** – An abnormal lab value is defined as one of the following:
  - Random glucose > 200 mg/dl
  - Fasting glucose ≥ 125 mg/dl
  - Hemoglobin A1c ≥ 6.5%

- **Physician entered diagnosis** – A physician entered diagnosis code is one that is derived from encounter or problem list sources only (excludes billing codes).

A flowchart expressing the logic of the T2DM case selection algorithm is shown in Figure 1. There are five paths through this flowchart, and these five paths are translated into equivalent pseudo-code in Algorithm 1. This algorithm shows the top-level logic, with additional sub-procedures implementing the lower-level details. These sub-procedures are also expressed below, in terms of an SQL-like syntax that is linked to the ICD-9, LOINC, and RxNorm codes in Appendix A.³

³These are the abbreviations used in the following flowcharts and algorithms:

- DX → DIAGNOSIS
- RX → PRESCRIPTION
- PT → PATIENT
- DT → DATE
- CNT → COUNT
Figure 1: Algorithm for identifying T2DM cases in the EMR.
Algorithm 1 T2DM case selection algorithm. This algorithm takes a patient-level record \((pt)\) as an argument, and returns the patient’s case status \((\{\text{CASE, UNKNOWN}\})\) as result.

\[
\begin{align*}
\text{T2DM-CASE-SELECTION}(pt) \\
\quad \text{status} &= \text{UNKNOWN} \\
1 \quad &\text{if } \text{T1DM-DX-DT-CNT}(pt) == 0 \\
&\quad \text{AND } \text{T2DM-DX-DT-CNT}(pt) > 0 \quad \text{\leftarrow Algorithm 2} \\
&\quad \text{AND } \text{T2DM-RX-DT}(pt) \neq \text{NULL} \quad \text{\leftarrow Algorithm 3} \\
&\quad \text{AND } \text{T1DM-RX-DT}(pt) \neq \text{NULL} \quad \text{\leftarrow Algorithm 4} \\
&\quad \text{AND } \text{T2DM-RX-DT}(pt) < \text{T1DM-RX-DT}(pt) \\
&\quad \quad \text{status} = \text{CASE} \\
2 \quad &\text{elseif } \text{T1DM-DX-DT-CNT}(pt) == 0 \\
&\quad \text{AND } \text{T2DM-DX-DT-CNT}(pt) > 0 \\
&\quad \text{AND } \text{T1DM-RX-DT}(pt) == \text{NULL} \\
&\quad \text{AND } \text{T2DM-RX-DT}(pt) \neq \text{NULL} \\
&\quad \quad \text{status} = \text{CASE} \\
3 \quad &\text{elseif } \text{T1DM-DX-DT-CNT}(pt) == 0 \\
&\quad \text{AND } \text{T2DM-DX-DT-CNT}(pt) > 0 \\
&\quad \text{AND } \text{T1DM-RX-DT}(pt) == \text{NULL} \\
&\quad \text{AND } \text{T2DM-RX-DT}(pt) == \text{NULL} \\
&\quad \text{AND } \text{ABNORMAL-LAB}(pt) == \text{TRUE} \quad \text{\leftarrow Algorithm 6} \\
&\quad \quad \text{status} = \text{CASE} \\
4 \quad &\text{elseif } \text{T1DM-DX-DT-CNT}(pt) == 0 \\
&\quad \text{AND } \text{T2DM-DX-DT-CNT}(pt) == 0 \\
&\quad \text{AND } \text{T2DM-RX-DT}(pt) \neq \text{NULL} \\
&\quad \text{AND } \text{ABNORMAL-LAB}(pt) == \text{TRUE} \\
&\quad \quad \text{status} = \text{CASE} \\
5 \quad &\text{elseif } \text{T1DM-DX-DT-CNT}(pt) == 0 \\
&\quad \text{AND } \text{T2DM-DX-DT-CNT}(pt) > 0 \\
&\quad \text{AND } \text{T1DM-RX-DT}(pt) \neq \text{NULL} \\
&\quad \text{AND } \text{T2DM-RX-DT}(pt) == \text{NULL} \\
&\quad \text{AND } \text{T2DM-PHYS CN-DX-DT-CNT}(pt) \geq 2 \quad \text{\leftarrow Algorithm 7} \\
&\quad \quad \text{status} = \text{CASE} \\
\text{return } \text{status}
\end{align*}
\]
**Algorithm 2** Count of distinct dates of T1DM DX (called by Algorithm 1)

\[T1DM-DX-DT-CNT(pt)\]

\[\text{count} = \]
\[\text{select COUNT-DISTINCT-DT(records)}\]
\[\text{from dx-table}\]
\[\text{where}\]
\[\text{dx-table.pt} == pt\]
\[\text{AND dx-table.icd-9-code} \in \{\ldots\} \iff \text{Table 3}\]
\[\text{return count}\]

**Algorithm 3** Count of distinct dates of T2DM DX (called by Algorithm 1)

\[T2DM-DX-DT-CNT(pt)\]

\[\text{count} = \]
\[\text{select COUNT-DISTINCT-DT(records)}\]
\[\text{from dx-table}\]
\[\text{where}\]
\[\text{dx-table.pt} == pt\]
\[\text{AND dx-table.icd-9-code} \in \{\ldots\} \iff \text{Table 4}\]
\[\text{return count}\]

**Algorithm 4** First date of Rx for T2DM medication (called by Algorithm 1)

\[T2DM-RX-DT(pt)\]

\[dt = \]
\[\text{select FIRST-DT(records)}\]
\[\text{from rx-table}\]
\[\text{where}\]
\[\text{rx-table.pt} == pt\]
\[\text{AND rx-table.rxnorm-code} \in \{\ldots\} \iff \text{Table 6}\]
\[\text{return dt}\]
**Algorithm 5** First date of Rx for T1DM medication (called by Algorithm 1)

```
T1DM-RX-DT(pt)

\[ dt = \] 
select FIRST-DT(records)
from rx-table
where
\[ rx-table.pt = pt \]
AND \[ rx-table.rxnorm-code \in \{\ldots\} \] \( \equiv \) Table 5

return \( dt \)
```

**Algorithm 6** Check for abnormal lab (called by Algorithm 1)

```
ABNORMAL-LAB(pt)

\[ abnormal-lab = \text{FALSE} \]
lab-results = 
select records
from labs-table
where
\[ labs-table.pt = pt \]
AND \[ labs-table.loinc-code \in \{\ldots\} \] \( \equiv \) Table 7

for each \( lab \in \text{lab-results} \)
\[ \text{if} \] \( lab.type = \text{RANDOM-GLUCOSE} \)
AND \( lab.value \geq 200 \) // (mg/dl)
OR \( lab.type = \text{FASTING-GLUCOSE} \)
AND \( lab.value \geq 125 \) // (mg/dl)
OR \( lab.type = \text{HBA1C} \)
AND \( lab.value \geq 6.5 \) // (percent)
\[ abnormal-lab = \text{TRUE} \]

return \( abnormal-lab \)
```
Algorithm 7 Count of distinct dates of T2DM DX made by a physician (called by Algorithm 1)

T2DM-PHYS-CN-DX-DT-CNT(\text{pt})

\[
\begin{align*}
\text{count} &= \text{select COUNT-DISTINCT-DT(\text{records})} \\
&\quad \text{from } \text{dx-table} \\
&\quad \text{where} \\
&\quad \quad \text{dx-table.pt} = \text{pt} \\
&\quad \quad \text{AND dx-table.source} \in \{\text{ENCOUNTER, PROBLEM-LIST}\} \\
&\quad \quad \text{AND dx-table.icd-9-code} \in \{\ldots\} \Leftarrow \text{Table 4} \\
\text{return count}
\end{align*}
\]
2.2 T2DM Control Selection Algorithm Logic

For the T2D control selection algorithm, the following data elements are required:

1. Counts of ICD-9 codes related to diabetes (Table 9)
2. Fasting blood glucose lab values – at least the maximum value (Table 7)
3. Random blood glucose lab values – at least the maximum value (Table 7)
4. HBA1c lab values – at least the maximum value (Table 7)
5. Diabetes family history – could be supplemented with self-reported data from a questionnaire
6. T1DM medications (i.e., Insulin & Symlin) order or prescription dates – at least the earliest date of Rx (Table 5)
7. T2DM medications order or prescription dates – at least the earliest date of Rx (Table 6)
8. Diabetes medical supply orders (Table 8)
9. Count of dates the patient had face-to-face outpatient clinic encounters (in-person visits with a clinician)

For the control selection algorithm, the following definition applies:

- **Abnormal lab**: abnormal lab values include:
  - Random glucose > 110 mg/dl
  - Fasting glucose ≥ 110 mg/dl
  - Hemoglobin A1c ≥ 6.0%

A flowchart expressing the logic of the T2DM control selection algorithm is shown in Figure 2. There is only one path through this flowchart, and this path is translated into equivalent pseudo-code in Algorithm 8. This algorithm shows the top-level logic, with additional sub-procedures implementing the lower-level details. These sub-procedures are also expressed below, in terms of an SQL-like syntax that is linked to the ICD-9, LOINC, and RxNorm codes in Appendix A.
Figure 2: Algorithm for identifying T2DM controls in the EMR.

**Algorithm 8** TT2DM control selection algorithm. This algorithm takes a patient-level record \( pt \) as an argument, and returns the patient’s control status \( \{ \text{CONTROL, UNKNOWN} \} \) as result.

**T2DM-CONTROL-SELECTION**(\( pt \))  
\[
\text{status} = \text{UNKNOWN}
\]

1. \( \text{if } \text{DM-DX-DT-CNT}(pt) = 0 \)  
2. AND GLUCOSE-LAB-EXISTS\( (pt) = \text{TRUE} \)  
3. AND ABNORMAL-LAB\( (pt) = \text{FALSE} \)  
4. AND ENCRTS-DT-CNT\( (pt) \geq 2 \)  
5. AND DM-MEDS-SUPPLIES-RX-DT-CNT\( (pt) = 0 \)  
6. AND FAM-HIST-OF-DM\( (pt) = \text{FALSE} \)  

\[
\text{return } \text{status} = \text{CONTROL}
\]
Algorithm 9 Count of distinct dates of DM-related DX (called by Algorithm 8)

DM-DX-DT-CNT($pt$)

\[
\text{count} = \\
\quad \text{select COUNT-DISTINCT-DT}(\text{records}) \\
\quad \text{from } dx\text{-table} \\
\quad \text{where} \\
\quad \quad dx\text{-table}.pt = pt \\
\quad \quad \text{AND } dx\text{-table.icd-9-code } \in \{\ldots\} \quad \Leftrightarrow \text{Table 9} \\
\text{return count}
\]

Algorithm 10 Check for glucose lab performed (called by Algorithm 8)

GLUCOSE-LAB-EXISTS($pt$)

\[
\text{glucose-lab-exists} = \text{false} \\
\text{lab-results} = \\
\quad \text{select records} \\
\quad \text{from } labs\text{-table} \\
\quad \text{where} \\
\quad \quad labs\text{-table}.pt = pt \\
\quad \quad \text{AND } labs\text{-table.loinc-code } \in \{\ldots\} \quad \Leftrightarrow \text{Table 7 (glucose only)} \\
\text{if } \text{lab-results.count} > 0 \\
\quad \text{glucose-lab-exists} = \text{true} \\
\text{return glucose-lab-exists}
\]
Algorithm 11 Check for abnormal lab (called by Algorithm 8)

ABNORMAL-LAB(pt)

\[
\text{abnormal-lab} = \text{FALSE}
\]

\[
\text{lab-results} =
\]

\[
\text{select records}
\]

\[
\text{from labs-table}
\]

\[
\text{where}
\]

\[
\text{labs-table.pt} == \text{pt}
\]

\[
\text{AND labs-table.loinc-code} \in \{ \ldots \} \quad \Leftarrow \text{Table 7}
\]

\[
\text{for each lab} \in \text{lab-results}
\]

\[
\text{if} \quad \text{lab.type} == \text{RANDOM-GLUCOSE}
\]

\[
\text{AND} \quad \text{lab.value} \geq 110 \quad \text{(mg/dl)}
\]

\[
\text{OR} \quad \text{lab.type} == \text{FASTING-GLUCOSE}
\]

\[
\text{AND} \quad \text{lab.value} \geq 110 \quad \text{(mg/dl)}
\]

\[
\text{OR} \quad \text{lab.type} == \text{HBA1C}
\]

\[
\text{AND} \quad \text{lab.value} \geq 6.0 \quad \text{(percent)}
\]

\[
\text{return} \quad \text{abnormal-lab} = \text{TRUE}
\]

Algorithm 12 Count of distinct dates for in-person office encounters with a physician (called by Algorithm 8)

ENCTRS-DT-CNT(pt)

\[
\text{count} =
\]

\[
\text{select COUNT-DISTINCT-DT(records)}
\]

\[
\text{from enctrs-table}
\]

\[
\text{where}
\]

\[
\text{enctrs-table.pt} == \text{pt}
\]

\[
\text{AND enctrs-table.type} == \text{OFFICE}
\]

\[
\text{return} \quad \text{count}
\]
**Algorithm 13** Count DM-related medications and supplies by distinct Rx date (called by Algorithm 8)

DM-MEDS-SUPPLIES-RX-DT-CNT(pt)
- \( count = \)
  - select COUNT-DISTINCT-DT(records)
  from rx-table
  where
    rx-table.pt == pt
  AND
    rx-table.rxnorm-code ∈ {...} ⇐ Table 5
  OR rx-table.rxnorm-code ∈ {...} ⇐ Table 6
  OR rx-table.rxnorm-code ∈ {...} ⇐ Table 8
- return count

**Algorithm 14** Check for family history of DM (called by Algorithm 8)

FAM-HIST-OF-DM(pt)
- fam-hist-results =
  - select records
  from fam-hist-table
  where
    fam-hist-table.pt == pt
  AND
    fam-hist-table.t1dm == TRUE
  OR fam-hist-table.t2dm == TRUE
- if fam-hist-results.count > 0
  return TRUE
- else return FALSE
3 KNIME workflow

This section describes installation of executable workflows that implement the case and control algorithms described in Section 2. These workflows are executed inside of the Konstanz Information Miner (KNIME) data analysis platform. The workflows take as input comma-separated value (csv) files, with each row corresponding to a patient (for examples, see the sample input files *dm_potential_cases.csv* and *dm_potential_controls.csv*).

3.1 Data Dictionaries

Each row of input data consists of a set of patient-level variables. We present here the data dictionaries that describe these patient-level input variables. The columns of the dictionaries specify each variable’s name, type, and range of possible values. The next column specifies whether or not missing values are permitted, and if so, what the default value of the variable is. The final column refers (where appropriate) to the table in Appendix A where corresponding code values for the variable can be found.

Table 1: Input variables to the T2DM case selection KNIME workflow

<table>
<thead>
<tr>
<th>Name</th>
<th>Type</th>
<th>Range</th>
<th>Missing (def.)</th>
<th>Ref.</th>
</tr>
</thead>
<tbody>
<tr>
<td>pat_id (unique)</td>
<td>INTEGER</td>
<td>( n \geq 1 )</td>
<td>FALSE</td>
<td>NA</td>
</tr>
<tr>
<td>t1dm_dx_cnt</td>
<td>INTEGER</td>
<td>( n \geq 0 )</td>
<td>TRUE (0)</td>
<td>Table 3</td>
</tr>
<tr>
<td>t2dm_dx_cnt</td>
<td>INTEGER</td>
<td>( n \geq 0 )</td>
<td>TRUE (0)</td>
<td>Table 4</td>
</tr>
<tr>
<td>t2dm_physcn_dx_cnt</td>
<td>INTEGER</td>
<td>( n \geq 0 )</td>
<td>TRUE (0)</td>
<td>Table 4</td>
</tr>
<tr>
<td>t1dm_rx_dt</td>
<td>STRING</td>
<td>yyyy-mm-dd</td>
<td>TRUE (NULL)</td>
<td>Table 5</td>
</tr>
<tr>
<td>t2dm_rx_dt</td>
<td>STRING</td>
<td>yyyy-mm-dd</td>
<td>TRUE (NULL)</td>
<td>Table 6</td>
</tr>
<tr>
<td>max_fast_gluc_lab_val</td>
<td>FLOAT</td>
<td>( n \geq 0.0 )</td>
<td>TRUE (NULL)</td>
<td>Table 7</td>
</tr>
<tr>
<td>max_rndm_gluc_lab_val</td>
<td>FLOAT</td>
<td>( n \geq 0.0 )</td>
<td>TRUE (NULL)</td>
<td>Table 7</td>
</tr>
<tr>
<td>max_hba1c_lab_val</td>
<td>FLOAT</td>
<td>( 0.0 \leq n \leq 100.0 )</td>
<td>TRUE (NULL)</td>
<td>Table 7</td>
</tr>
</tbody>
</table>

The default value is automatically inserted for a variable when it is missing a specified value.
<table>
<thead>
<tr>
<th>Name</th>
<th>Type</th>
<th>Range</th>
<th>Missing (def.)</th>
<th>Ref.</th>
</tr>
</thead>
<tbody>
<tr>
<td>pat_id (unique)</td>
<td>INTEGER</td>
<td>$n \geq 1$</td>
<td>FALSE</td>
<td>NA</td>
</tr>
<tr>
<td>fam_hist_of_dm</td>
<td>INTEGER</td>
<td>$n \in {0, 1}$</td>
<td>TRUE (0)</td>
<td>NA</td>
</tr>
<tr>
<td>encnrs_cnt</td>
<td>INTEGER</td>
<td>$n \geq 0$</td>
<td>TRUE (0)</td>
<td>Table 3</td>
</tr>
<tr>
<td>max_fast_gluc_lab_val</td>
<td>FLOAT</td>
<td>$n \geq 0.0$</td>
<td>TRUE (NULL)</td>
<td>Table 7</td>
</tr>
<tr>
<td>max_rndm_gluc_lab_val</td>
<td>FLOAT</td>
<td>$n \geq 0.0$</td>
<td>TRUE (NULL)</td>
<td>Table 7</td>
</tr>
<tr>
<td>max_hba1c_lab_val</td>
<td>FLOAT (%)</td>
<td>$0.0 \leq n \leq 100.0$</td>
<td>TRUE (NULL)</td>
<td>Table 7</td>
</tr>
<tr>
<td>dm_dx_cnt</td>
<td>INTEGER</td>
<td>$n \geq 0$</td>
<td>TRUE (0)</td>
<td>Table 9</td>
</tr>
<tr>
<td>dm_med_supplies_cnt</td>
<td>INTEGER</td>
<td>$n \geq 0$</td>
<td>TRUE (0)</td>
<td>Table 5,</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Table 6,</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Table 8</td>
</tr>
</tbody>
</table>
3.2 Installation and Execution

1. Download and install KNIME (version 2.4 or later). The KNIME website contains installation instructions, as well as tutorials.

2. Download the T2D case and control workflows, which are contained in a single zip file: T2D-workflows.zip. Don’t unzip the file.

3. Download the two sample input files for the workflows: dm_potential_cases.csv and dm_potential_controls.csv.

4. Start KNIME. On start-up, you will see an empty workspace similar to the screenshot in Figure 3.

5. Select File ⇒ Import KNIME workflow. The resulting pop-up window is shown in Figure 4. Click on the Select archive file: radio button, and navigate to your local copy of the T2D-workflows.zip file. Click on the Finish button.

6. Double-click on the Diabetes-Case-Assignment workflow to open it. Your workspace will now look similar to the screenshot in Figure 5.

7. Double-click on the File Reader node in the workflow graph. You will see the pop-up window shown in Figure 6. Click on the Browse... button and navigate to your local copy of the dm_potential_cases.csv file. Your pop-up window should look like the one in Figure 7. Make sure that the read row IDs box is unchecked, while the read column headers box is checked. Click on the OK button to close the window.

8. Double-click on the CSV Writer node in the workflow graph. Click on the Browse... button and navigate to a directory of your choosing where the output file dm_cases.csv will be generated. Click on the OK button to close the window.

9. The workflow is now ready to execute. Click on the green button with the double arrow at the toolbar at the top, or enter Shift+F7 on the keyboard. If the nodes of the workflow have already been executed, then first select all nodes (Control+A), right click, and select Reset.

   (a) The output file will be located in the directory that you chose in Step 8.

   (b) Right click on the Rule Engine node in the graphical view and select Classified Data in order to view the output of the algorithm (Figure 8).

   (c) Right click on the Histogram node in the graphical view and select View:Histogram View to get counts of the assignments that were made (Figure 9).

---

5 All following steps apply also to the Diabetes-Control-Assignment workflow.
6 An executed node will have a green indicator underneath it.
Figure 3: Step 4
Figure 4: Step 5

![Workflow Import Selection](image)

- **Source:**
  - Select root directory:
  - Select archive file:

- **Target:**
  - Select destination:

- **Workflows:**
  - Diabetes-Cases Assignment
  - Diabetes-Control Assignment

- Copy projects into workspace

- Buttons: < Back, Next, Finish, Cancel
Figure 5: Step 6
Figure 6: Step 7

Figure 7: Step 7
Figure 8: Step 9b

Figure 9: Step 9c
## A Data Elements

Table 3: T1DM diagnosis codes. Used in Algorithm 1.

<table>
<thead>
<tr>
<th>Description</th>
<th>ICD-9 code</th>
</tr>
</thead>
<tbody>
<tr>
<td>Type 1 Diabetes</td>
<td>250.x1, 250.x3</td>
</tr>
</tbody>
</table>

Table 4: T2DM diagnosis codes. Used in Algorithm 1.

<table>
<thead>
<tr>
<th>Description</th>
<th>ICD-9 code</th>
</tr>
</thead>
<tbody>
<tr>
<td>Type 2 Diabetes</td>
<td>250.x0, 250.x2 (excl. 250.10, 250.12)</td>
</tr>
</tbody>
</table>

Table 5: T1DM medications. Used in Algorithm 1 and Algorithm 8.

<table>
<thead>
<tr>
<th>Generic Name</th>
<th>Example Brand</th>
<th>RxNorm CUI (ingredient-level)</th>
</tr>
</thead>
<tbody>
<tr>
<td>insulin</td>
<td></td>
<td>139825, 274783, 314684, 352385, 400008, 51428, 5856, 86009</td>
</tr>
<tr>
<td>pramlintide</td>
<td>Symlin</td>
<td>139953</td>
</tr>
</tbody>
</table>
Table 6: T2DM medications. Used in Algorithm 1 and Algorithm 8.

<table>
<thead>
<tr>
<th>Generic Name</th>
<th>Example Brand</th>
<th>RxNorm CUI</th>
</tr>
</thead>
<tbody>
<tr>
<td>acetohexamide</td>
<td>Dymelor</td>
<td>173</td>
</tr>
<tr>
<td>tolazamide</td>
<td>Tolinase</td>
<td>10633</td>
</tr>
<tr>
<td>chlorpropamide</td>
<td>Diabinese</td>
<td>2404</td>
</tr>
<tr>
<td>glipizide</td>
<td>Glucotrol</td>
<td>4821</td>
</tr>
<tr>
<td>glimepiride</td>
<td>Glucotrol XL</td>
<td>217360</td>
</tr>
<tr>
<td>glyburide</td>
<td>Micronase, Glynase, Diabeta</td>
<td>4815</td>
</tr>
<tr>
<td>glimepiride</td>
<td>Amaryl</td>
<td>25789</td>
</tr>
<tr>
<td>repaglinide</td>
<td>Prandin</td>
<td>73044</td>
</tr>
<tr>
<td>nateglinide</td>
<td>Starlix</td>
<td>274332</td>
</tr>
<tr>
<td>metformin</td>
<td>Glucophage</td>
<td>6809</td>
</tr>
<tr>
<td>rosiglitazone</td>
<td>Avandia</td>
<td>84108</td>
</tr>
<tr>
<td>pioglitazone</td>
<td>ACTOS</td>
<td>33738</td>
</tr>
<tr>
<td>troglitazone</td>
<td>Rezulin</td>
<td>72610</td>
</tr>
<tr>
<td>acarbose</td>
<td>Precose</td>
<td>16681</td>
</tr>
<tr>
<td>miglitol</td>
<td>Glyset</td>
<td>30009</td>
</tr>
<tr>
<td>sitagliptin</td>
<td>Januvia</td>
<td>593411</td>
</tr>
<tr>
<td>exenatide</td>
<td>Byetta</td>
<td>60548</td>
</tr>
</tbody>
</table>

Table 7: Diabetes mellitus lab codes. Used in Algorithm 1 and Algorithm 8

<table>
<thead>
<tr>
<th>Description</th>
<th>LOINC code</th>
</tr>
</thead>
<tbody>
<tr>
<td>Fasting glucose</td>
<td>1558-6</td>
</tr>
<tr>
<td>Random glucose</td>
<td>2339-0, 2345-7</td>
</tr>
<tr>
<td>Hemoglobin A1C</td>
<td>4548-4, 17856-6, 4549-2, 17855-8</td>
</tr>
</tbody>
</table>
Table 8: Diabetes medical supplies. Used in Algorithm 8.

<table>
<thead>
<tr>
<th>Description</th>
<th>Source Vocab.</th>
<th>RxNorm CUI (ingredient-level)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Blood-glucose meters &amp; sensors</td>
<td>NDDF</td>
<td>126958, 412956, 412959, 637321, 668291, 668370, 686655, 692383, 748611, 880998, 881056</td>
</tr>
<tr>
<td></td>
<td>VANDF</td>
<td>751128</td>
</tr>
<tr>
<td>Insulin syringes</td>
<td>RxNorm</td>
<td>847187, 847191, 847197, 847203, 847207, 847211, 847230, 847239, 847252, 847256, 847259, 847263, 847278, 847416, 847417</td>
</tr>
<tr>
<td></td>
<td>NDDF</td>
<td>806905, 806903, 408119</td>
</tr>
</tbody>
</table>

Table 9: Diabetes mellitus diagnosis codes. Used in Algorithm 8.

<table>
<thead>
<tr>
<th>Description</th>
<th>ICD-9 code</th>
</tr>
</thead>
<tbody>
<tr>
<td>Diabetes mellitus (T1 &amp; T2)</td>
<td>250.xx</td>
</tr>
<tr>
<td>Impaired fasting glucose</td>
<td>790.21</td>
</tr>
<tr>
<td>Impaired oral glucose tolerance test</td>
<td>790.22</td>
</tr>
<tr>
<td>Abnormal glucose not otherwise spec.</td>
<td>790.2, 790.29</td>
</tr>
<tr>
<td>Abnormal glucose during pregnancy</td>
<td>648.8x</td>
</tr>
<tr>
<td>Gestational diabetes</td>
<td>648.0x</td>
</tr>
<tr>
<td>Glycosuria</td>
<td>791.5</td>
</tr>
<tr>
<td>Dysmetabolic syndrome X</td>
<td>277.7</td>
</tr>
<tr>
<td>Family history of diabetes mellitus</td>
<td>V18.0</td>
</tr>
<tr>
<td>Screening for diabetes mellitus</td>
<td>V77.1</td>
</tr>
</tbody>
</table>