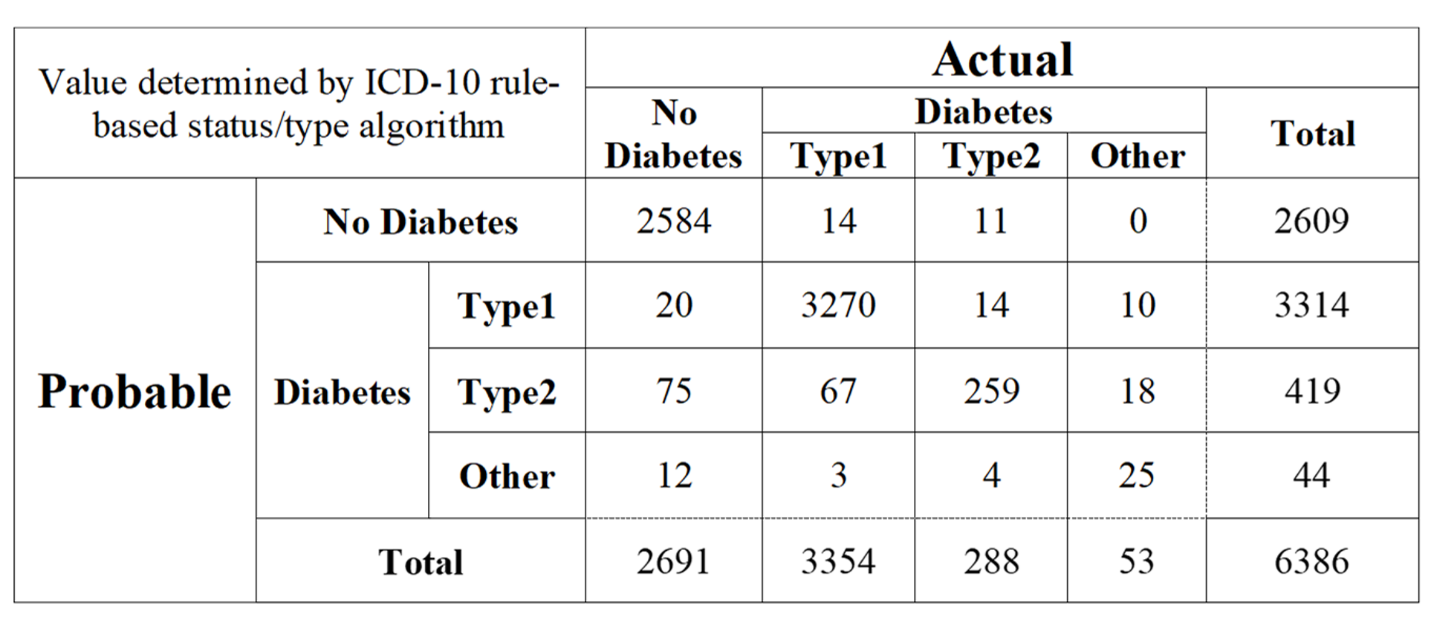
**Supplemental Material**

**Supplemental Table 1.** Time in EHR for diabetes cases identified by the rule-based ICD-10 status and type algorithm.

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| Rule-based ICD-10 Algorithm Diabetes Type | **Type 1** | **Type 2** | **Other** | **Total** |
| Case Count | 3314 | 419 | 44 | 3777 |
| Mean total months in EHR  (SD) | 60.5 (30.3) | 57.2 (32.6) | 51.2 (32.5) | 60.0 (30.6) |
| Diagnosis calendar month/year is same as EHR entry  n (%) | 1799 (54.3) | 111 (26.5) | 6 (13.6) | 1916 (50.7) |
| **Evidence: Time occurrence of two or more ICD diabetes codes** | | | | |
| Mean months without evidence in months (SD) | 14.1 (24.6) | 32.70 (32.13) | 26.11 (28.19) | 16.24 (26.22) |
| Enter EHR and evidence within the same month  n (%) | 2114 (63.8) | 133 (31.7) | 8 (18.2) | 2255 (59.7) |
| Entry into EHR and first diabetes evidence within 6 months  n (%) | 2258 (68.1) | 161 (38.4) | 16 (36.4) | 2435 (64.5) |
| Entry into EHR and first diabetes evidence within 12 months  n (%) | 2357 (71.1) | 177 (42.2) | 18 (40.9) | 2552 (67.6) |
| **Evidence: First occurrence of any item in the multiple-criteria algorithm** | | | | |
| Mean months without evidence in months (SD) | 13.7 (24.4) | 26.6 (29.9) | 18.89 (23.0) | 15.2 (25.4) |
| Enter EHR and evidence within the same month  n (%) | 2165 (65.3) | 152 (36.3) | 14 (31.8) | 2331 (61.7) |
| Entry into EHR and first diabetes evidence within 6 months  n (%) | 2281 (68.8) | 178 (42.5) | 20 (45.5) | 2479 (65.6) |
| Entry into EHR and first diabetes evidence within 12 months  n (%) | 2379 (71.8) | 195 (46.5) | 23 (52.3) | 2597 (68.8) |

Start of time in the EHR was considered the first month and year in which a patient had any recorded EHR measurement within any domain (diagnosis codes, vitals, laboratory measurements, and medications.)

**Supplemental Figure 1**: Diabetes status and type classification matrix of rule-based ICD-10 status/type algorithm.

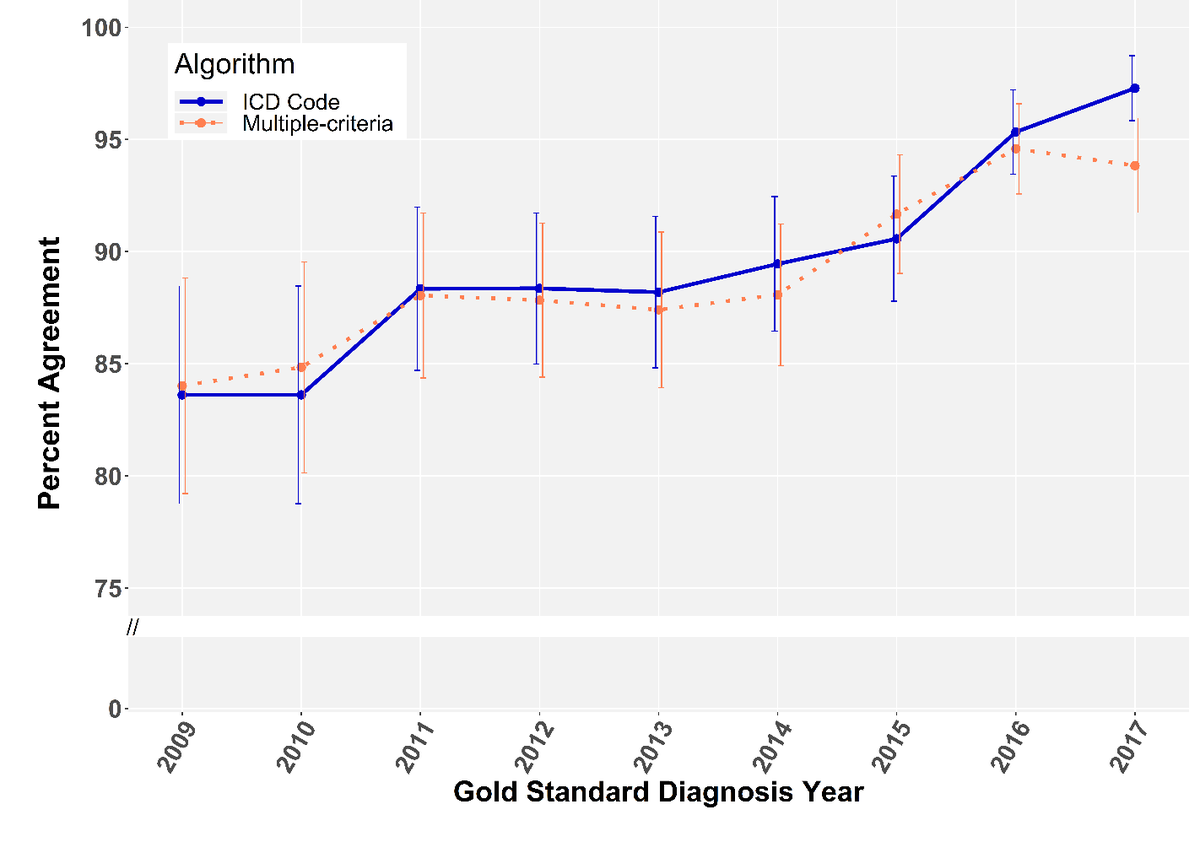


This ICD-10 algorithm identified a sub-population of cases within which date of diagnosis algorithms were tested. This classification matrix shows comparison to gold standard status and type identified.

**Supplemental Figure 2:** Algorithm percent agreement with gold standard year of diagnosis by sit

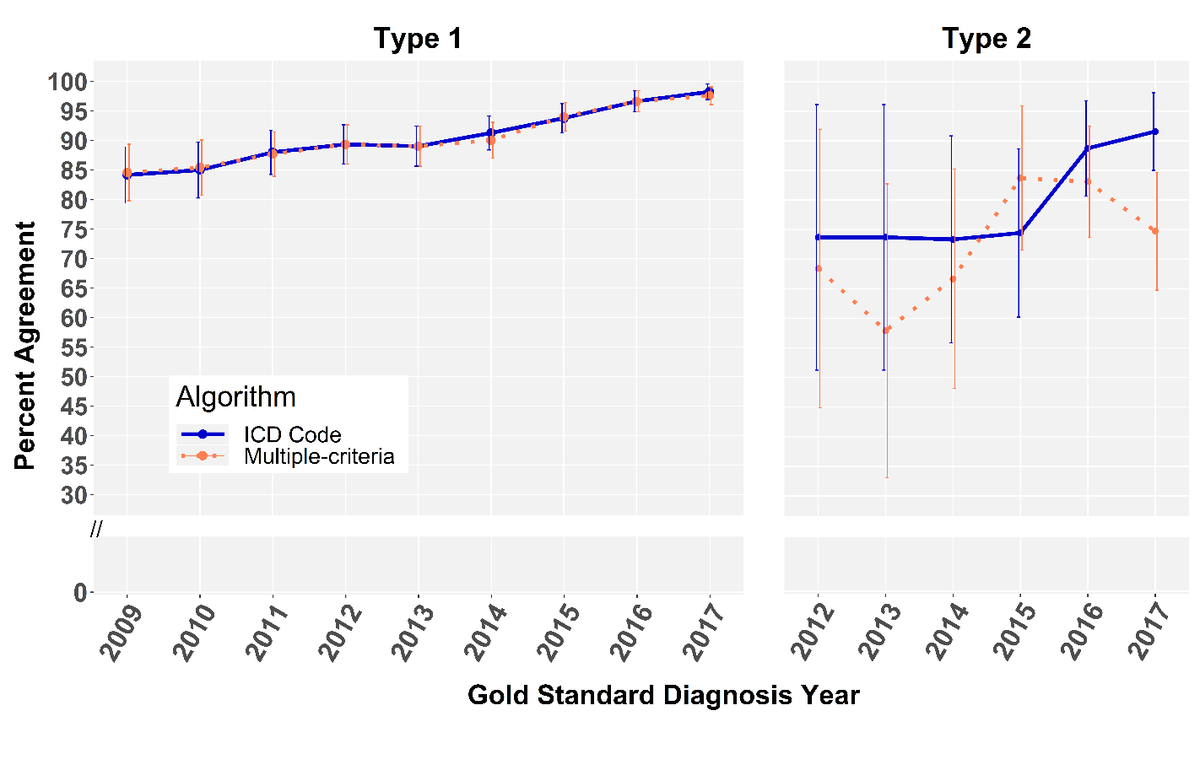
Non-diabetes observations (n=107) incorrectly identified by the rule-based ICD-10 algorithm and diabetes cases (n=119) with gold standard date of diagnosis preceding 2009 are not visualized.

**Supplemental Figure 3**: Algorithm percent agreement with calendar month/year (+ One Month).



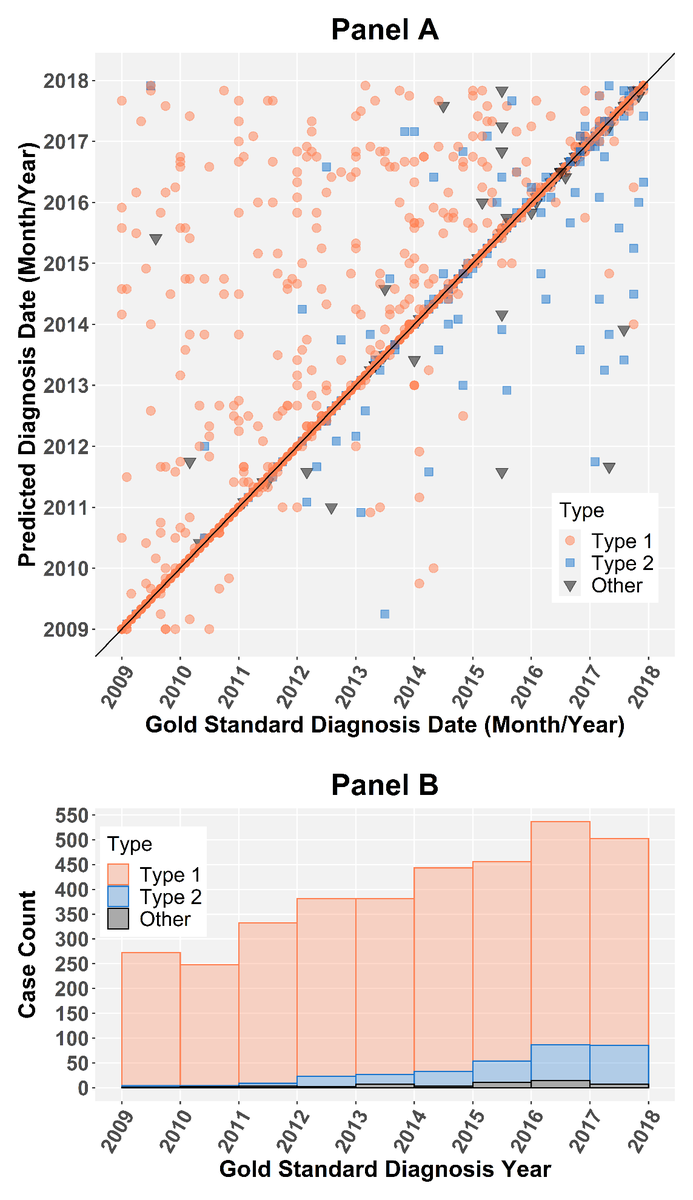
Results aggregated by diagnosis year. Non-diabetes observations (n=107) incorrectly identified by the ICD-10 rule-based algorithm and diabetes cases (n=119) with gold standard date of diagnosis preceding 2009 are not visualized.

**Supplemental Figure 4:** Algorithm percent agreement with calendar month/year (+ One Month).



Results aggregated by diagnosis year. Non-diabetes observations (n=107) incorrectly identified by the ICD-10 rule-based algorithm and diabetes cases (n=119) with gold standard date of diagnosis preceding 2009 are not visualized.

**Supplemental Figure 5**: Multiple-criteria algorithm and gold standard date of diagnosis concordance.



In Panel A, the diagonal line represents perfect alignment of calendar month/year between the predicted and gold standard date of diagnosis for diabetes cases. The accompanying histogram in Panel B demonstrates frequency of diabetes cases by type and within each year as the intensity of the 45 degree line is not easily discernable.