

Supplemental Table S1. Details of the Wide Net algorithm.

Individuals are considered to have possible diabetes if they meet any of the following criteria
<ul style="list-style-type: none"> • Hemoglobin A1c $\geq 6.5\%$ (≥ 48 mmol/mol)
OR
<ul style="list-style-type: none"> • Fasting plasma glucose ≥ 126mg/dl (≥ 7.0 mmol/L)
OR
<ul style="list-style-type: none"> • Random plasma glucose ≥ 200 mg/dl (≥ 11.1 mmol/L)
OR
<ul style="list-style-type: none"> • At least one of the following diabetes-related ICD-10 encounter diagnoses codes: E08, E09, E10, E11, or E13.
OR
<ul style="list-style-type: none"> • A prescription for or administration of one or more medications residing in one of the following antidiabetic medication classes*:
<ul style="list-style-type: none"> <ul style="list-style-type: none"> ○ Alpha Glucosidase Inhibitors
<ul style="list-style-type: none"> <ul style="list-style-type: none"> ○ Dipeptidyl Peptidase-4 (DPP4) Inhibitors
<ul style="list-style-type: none"> <ul style="list-style-type: none"> ○ Glucagon-like Protein-1 (GLP-1) Receptor Agonists
<ul style="list-style-type: none"> <ul style="list-style-type: none"> ○ Insulin
<ul style="list-style-type: none"> <ul style="list-style-type: none"> ○ Meglitinides
<ul style="list-style-type: none"> <ul style="list-style-type: none"> ○ Sodium-glucose co-transporter-2 (SGLT2) inhibitors
<ul style="list-style-type: none"> <ul style="list-style-type: none"> ○ Sulfonylureas
<ul style="list-style-type: none"> <ul style="list-style-type: none"> ○ Thiazolidinediones
<ul style="list-style-type: none"> <ul style="list-style-type: none"> ○ Other

* Lists of RxNorm concept unique identifiers (CUIs) for specific medications occurring in each of these classes is included as a .csv file in the supplemental material

Supplemental Table S2. Candidate variables defined

Description	Definition	Included in Final Regression Model
Maximum HbA1c	highest A1c value observed before 1/1/2018. Variable was categorized into one of the following levels: <5.7% (39mmol/mol, 5.7-6.4% (39mmol/mol-48mmol/mol, >=6.5% (48mmol/mol), not measured	X
Antihyperlipidemics†	Patient was ever prescribed an antihyperlipidemic medication prior to 1/1/2018. (Yes or No)	
Outpatient antihypertensive†	Patient was ever prescribed an antihypertensive medication prior to 1/1/2018. (Yes or No)	
Any antidiabetic medicine†	Patient was ever prescribed an antidiabetic medication prior to 1/1/2018. (Yes or No)	X
Any diabetic ketoacidosis*	Any ICD-10 encounter diagnosis for DKA prior to 1/1/2018	X
Ethnicity	Hispanic or non-Hispanic	X
Any elevated outpatient glucose	Any elevated values of outpatient glucose (random >=200 mg/dl, fasting >=126 mg/dl) via venipuncture or fingerstick prior to 1/1/2018. (Yes or No)	X
Multiple elevations in outpatient random glucose	Elevated values of outpatient random glucose (>=200 mg/dl) via venipuncture or fingerstick prior to 1/1/2018 on at least 2 separate occasions.	X
Median high-density lipoprotein	Median of all serum HDL values (mg/dl) obtained prior to 1/1/2018, categorized as: 1 st quartile, 2 nd quartile, 3 rd quartile, 4 th quartile, or missing.	
Median low-density lipoprotein	Median of all serum LDL values (mg/dl) obtained prior to 1/1/2018, categorized as: 1 st quartile, 2 nd quartile, 3 rd quartile, 4 th quartile, or missing.	
Any contraceptive medication†	Patient was ever prescribed a contraceptive medication prior to 12/31/2018. (Yes or No)	X
Any diabetes*	Any ICD-10 encounter diagnosis for diabetes (all types) prior to 1/1/2018. (Yes or No)	
Hyperlipidemia*	Any ICD-10 encounter diagnosis for hyperlipidemia prior to 1/1/2018. (Yes or No)	
Hypertension*	Any ICD-10 encounter diagnosis for hypertension prior to 1/1/2018. (Yes or No)	
Menstrual irregularities*	Any ICD-10 encounter diagnosis for menstrual irregularities prior to 1/1/2018. (Yes or No)	
Obesity*	Any ICD-10 encounter diagnosis for obesity prior to 1/1/2018. (Yes or No)	X
Other diabetes type*	Any ICD-10 diabetes specific encounter diagnoses codes for diabetes types “other” than type 1 or type 2 prior to 1/1/2018. (Yes or No)	
PCOS*	Any ICD-10 encounter diagnosis for PCOS prior to 1/1/2018. (Yes or No)	X
Race	White, black, or other	
Gender	Male or female	
Median triglyceride	Median of all outpatient serum triglyceride values (mg/dl) obtained prior to 1/1/2018, categorized as: 1 st quartile, 2 nd quartile, 3 rd quartile, 4 th quartile, or missing.	
Any metformin†	Patient was ever prescribed metformin prior to 12/31/2018. (Yes or No)	X
Proportion of “other” diabetes codes	Count of “other” diabetes type ICD-10 codes / total count of all diabetes codes prior to 1/1/2018. Variable was categorized into one of the following levels: 0-0.24, 0.25-0.49, 0.5-0.74, 0.75-1	X
Proportion of type 1 diabetes codes	Count of type 1 specific diabetes ICD-10 codes / total count of all diabetes codes prior to 1/1/2018. Variable was categorized into one of the following levels: 0-0.24, 0.25-0.49, 0.5-0.74, 0.75-1	X
Proportion of type 2 diabetes codes	Count of type 2 specific diabetes ICD-10 codes / total count of all diabetes codes prior to 1/1/2018. Variable was categorized into one of the following levels: 0-0.24, 0.25-0.49, 0.5-0.74, 0.75-1	X

Most common diabetes type specific code	Which diabetes type specific ICD 10 encounter diagnoses occurred the most frequently. Categorized as: <2 codes, type1, type 2, other	X
Age	Age in years as of 12/31/2017. Modelled as a continuous variable using restricted cubic splines with 3 knots.	X
Count of “other” diabetes codes	number of times that an “other” diabetes type ICD 10 code occurred before 1/1/2018. Modelled as a continuous variable using restricted cubic splines with 3 knots.	
Count of type 1 diabetes codes	number of times that type 1 diabetes ICD 10 code occurred before 1/1/2018. Modelled as a continuous variable using restricted cubic splines with 3 knots.	X
Count of type 2 diabetes codes	number of times that an “other” diabetes type ICD 10 code occurred before 1/1/2018. Modelled as a continuous variable using restricted cubic splines with 3 knots.	

* Diagnoses were dichotomous variables based on the presence of one or more structured diagnoses code(s) (ICD10) prior to 1/1/2018. Specific ICD10 codes provided in supplemental table S4.

† list of specific generic medication names and associated RxNorm codes are provided in the supplemental medication comma separated value file (.csv).

PCOS = Polycystic Ovarian Syndrome

DKA = Diabetic Keto-Acidosis

HDL = High density lipoprotein

LDL = Low density lipoprotein

Supplemental Table S3. Performance of multinomial regression and a rule-based algorithm using ICD 10 codes for determining diabetes status and type by site using SEARCH cohort status as gold standard.

		Multinomial Regression			Rule-Based Algorithm		
	Site	Ohio*	Washington†	Colorado†	Ohio	Washington	Colorado
	N	3742	2260	2680	3742	2260	2680
Diabetes n = 5,308	<i>Se</i>	0.985	0.922	0.985	0.988	0.988	0.997
	<i>Sp</i>	0.978	0.993	0.976	0.972	0.972	0.954
	PPV	0.970	0.997	0.992	0.932	0.990	0.985
	NPV	0.989	0.813	0.954	0.992	0.966	0.990
Type 1 Diabetes n = 4,732	<i>Se</i>	0.891	0.981	0.986	0.975	0.968	0.991
	<i>Sp</i>	0.989	0.969	0.931	0.988	0.986	0.928
	PPV	0.980	0.985	0.968	0.980	0.993	0.967
	NPV	0.939	0.961	0.969	0.985	0.938	0.979
Type 2 Diabetes n = 400	<i>Se</i>	0.692	0.172	0.856	0.899	0.910	0.887
	<i>Sp</i>	0.991	0.998	0.988	0.965	0.975	0.985
	PPV	0.791	0.821	0.722	0.547	0.697	0.683
	NPV	0.986	0.950	0.995	0.995	0.994	0.996

Other Diabetes Type n = 176 (e.g., medication- induced, monogenic)	<i>Se</i>	0.679	0.240	0.224	0.642	0.480	0.367
	<i>Sp</i>	0.949	0.999	0.997	0.996	0.997	0.995
	PPV	0.161	0.667	0.710	0.694	0.667	0.735
	NPV	0.995	0.992	0.971	0.995	0.994	0.976
Kappa Statistic		0.870	0.846	0.894	0.909	0.919	0.903
Accuracy		0.928	0.928	0.951	0.950	0.961	0.955

Se = Sensitivity

Sp = Specificity

PPV = Positive Predictive Value

NPV = Negative Predictive Value

LR+ = Positive Likelihood Ratio

LR- = Negative Likelihood Ratio

Accuracy = number correctly classified / N

Positive and Negative Likelihood Ratios may be calculated with the following formulas:

$LR+ = Se/(1 - Sp)$, $LR- = (1 - Se)/Sp$

* Cincinnati Children's Hospital

† Colorado Children's Hospital

‡ Seattle Children's Hospital

Supplemental Table S4. ICD-10 codes used for each condition

Diagnosis	ICD-10 codes
Hypertension	I10-I15.x
Hyperlipidemia	E78.0-E78.5.x
Obesity	E66.x
PCOS	E28.2x
Menstrual Irregularities	N91-N92.x, N93.1, N93.8, N93.9
Diabetes	E08-E13.x, P70.2, O24.0x, O24.1x