Supplemental material

BMJ Publishing Group Limited (BMJ) disclaims all liability and responsibility arising from any reliance placed on this supplemental material which has been supplied by the author(s)

BMJ Open

Online Supplement

Table S1. Diabetes case finding using EHR data.

Must meet at least one of the following criteria:

1. At least two separate encounter dates (inpatient, outpatient, emergency department) with type 2 diabetes diagnosis codes (ICD-9, ICD-10, or electronic diagnosis group [EDG]).
   a. Excluded if had ≥ 10 years of type 1 diagnoses and < five years with type 2 diagnoses.
   b. Excluded if < 10 years of age at first diabetes diagnosis.
2. At least one diabetes medication order, other than metformin or acarbose if female. Metformin combination medications were included.
   a. Excluded if first diabetes medication order was prior to age 10 years.
3. At least one encounter with type 2 diabetes diagnosis and an abnormal laboratory value (random glucose ≥ 200 mg/dl; fasting glucose ≥ 126 mg/dl; or hemoglobin A1c ≥ 6.5%).
   a. Excluded if had ≥ 10 years of type 1 diagnoses and < five years with type 2 diagnoses.

- The date of onset was assigned as the earliest date with any evidence of diabetes (e.g., had generic diabetes diagnoses that were not used for definition #1, or had abnormal laboratory value that was not accompanied by a diagnosis so did not meet definition #3).

Notes:
a) To meet criteria #2 or #3, criterion had to occur > 9 months prior to or > 1 month after delivery of child (to avoid gestational diabetes). Gestational diabetes was not an exclusion if the individual subsequently developed type 2 diabetes. Date of onset was assigned as when the person met the type 2 diabetes criterion; and
b) EDG codes are used in Epic EHR software (Epic Systems Corporation, Verona, WI) and often have higher specificity and greater detail.
c) Of the 15,888 diabetes cases: 11,944 met criterion 1; 10,183 met criterion 2; 12,552 met criterion 3; 7008 met all three; and 4775 met at least two.
d) Because metformin can be used for pre-diabetes, we evaluated how many persons could have had this diagnosis instead of diabetes in our diabetes onset definition. Of the 1579 men who met only definition #2, between 544 (3.4%) and 1207 (7.6%) may have had pre-diabetes instead of diabetes, depending on how longitudinal information on diagnoses, medications, medication indications, and abnormal laboratory results were used and interpreted.

Table S2. Selected characteristics of study individuals and communities by administrative community type.

Variables	Borough	Census Tract	Township
By community type (n = 1070 communities)			
Number (%), total	291 (27.2)	146 (13.6)	633 (59.2)
Number (%), among cases	224 (27.6)	107 (13.2)	482 (59.3)
Number (%), among controls	278 (26.9)	137 (13.2)	620 (59.9)
Counties with at least one resident in community type, n	35	16	37
Counties with at least 20 residents in community type, n	27	9	32
Community measures, by community type (n = 1070 communities)			
Area, square miles, mean (SD)	1.72 (2.32)	1.20 (3.52)	29.4 (18.1)
Community socioeconomic deprivation, mean (SD)	-0.09 (2.99)	4.17 (3.80)	-1.15 (2.71)
Population density, persons per square mile, mean (SD)	2094.7 (1642.3)	6594.5 (5014.6)	157.5 (279.4)
Developed land, % (SD)	37.2 (22.6)	72.6 (23.0)	3.66 (7.35)
Intersection density per square mile, mean (SD)	120.6 (86.1)	208.5 (117.0)	13.34 (14.77)
By participant (n = 95,323 individuals)			
Cases, n (%) (total = 15,888)	4621 (29.1)	1806 (11.4)	9461 (59.5)
Controls, n (%) (total = 79,435)	21,756 (27.4)	6548 (8.2)	51,131 (64.4)
Age at diabetes onset or control selection date, years, mean (SD)	54.4 (15.9)	52.7 (16.1)	55.3 (14.8)
Sex, female, n (%)	13,329 (50.2)	4449 (53.3)	29,098 (48.0)
Race, white, n (%)	245,963 (98.4)	7873 (94.2)	59,460 (98.1)
Ethnicity, Hispanic, n (%)	353 (1.3)	430 (5.2)	680 (1.1)
Body mass index, kg/m^2, mean (SD)	30.6 (7.47)	30.9 (7.96)	30.3 (6.94)
Medical Assistance, % of time, mean (SD)	5.9 (17.9)	10.3 (23.2)	3.3 (13.5)
Medical Assistance, ever*, n (%)	3311 (12.6)	1692 (20.3)	4311 (7.1)
Contact with health system <u>before</u> diagnosis/control selection date, years, mean (SD)	12.7 (4.37)	12.1 (4.57)	12.9 (4.34)
Charlson index, mean (SD)	1.75 (1.83)	1.64 (1.78)	1.76 (1.78)
Greenness, peak NDVI, in buffer, mean (SD)	0.61 (0.11)	0.51 (0.10)	0.73 (0.10)
Urban status by UA and UC boundaries, n (col %)			
Rural	3031 (11.5)	10 (0.1)	38,456 (63.5)
Urbanized area (UA)	11,409 (43.3)	5414 (64.8)	11,506 (19.0)
Urban cluster (UC)	11,937 (45.3)	2930 (35.1)	10,630 (17.5)
Abbreviations: NDVI = normalized difference vegetation index; SD = standard deviation. ** At least one encounter that used Medical Assistance for health insurance.*			

Table S3. Mean outpatient encounters among cases and controls by community type and Medical Assistance status.

Variable	Cases, n = 15,888			Controls, n = 79,435		
	Boroughs n = 4621	City Census Tracts n = 1806	Townships n = 9461	Boroughs n = 21,756	City Census Tracts n = 6548	Townships n = 51,131
Outpatient encounters, total before diagnosis, mean (SD)	35.9 (34.8)	31.6 (32.1)	36.8 (35.2)	35.7 (33.8)	33.5 (32.8)	35.2 (31.8)
Outpatient encounters before diagnosis, mean (SD), <u>by Medical Assistance status</u> (% time receiving)						
0%	35.2 (33.9)	30.9 (31.2)	36.3 (34.6)	35.1 (33.1)	32.9 (32.1)	35.0 (31.7)
0.1-24.9%	47.7 (41.3)	41.8 (44.3)	44.7 (39.0)	44.2 (40.7)	40.0 (39.9)	42.6 (36.1)
25.0-74.9%	32.5 (34.5)	29.3 (25.4)	37.1 (40.2)	37.3 (36.8)	33.6 (32.4)	34.2 (31.4)
75+%	30.6 (28.9)	34.2 (21.0)	30.7 (28.0)	27.7 (28.5)	27.9 (28.4)	27.7 (22.6)
SD = standard deviation						

3

Medical Profile of Cases and Controls

To evaluate our categorization of diabetes cases and controls, we examined a number of biomarkers and other measures of relevance to diabetes, dysglycemia, and other cardio-metabolic risk factors development that were available in the EHR, including hemoglobin A1c (HbA1c), lipids (cholesterol and triglycerides), blood glucose (fasting and unspecified), and body mass index (BMI) (**Online Supplement Table S4**). Fasting blood glucose was measured in the year before the diabetes onset or control dates in 24% of cases and 29% of controls. Interestingly, the mean value was higher in the year before diagnosis in persons who would develop diabetes compared to those who would not, 108.5 vs. 95.8 mg/dL ($p < 0.001$). In the year after diagnosis or control dates, fasting blood glucose was available in 58% of cases and 30% of controls, and mean levels were much higher in cases compared to controls (147.9 vs. 95.9, $p < 0.001$). HbA1c, triglycerides, unspecified blood glucose, and BMI all evidenced similar patterns (**Online Supplement Table S4**). In the year before and after diagnosis, most cases and controls had BMI measured, with a much higher mean in cases compared to controls before and after diagnosis.

Table S4. Selected laboratory and other biometric values comparing new onset type 2 diabetes cases and controls without diabetes.

Variable	Cases	Controls
Number	15,888	79,435
Hemoglobin A1c (HbA1c)		
# in year <u>before</u> diagnosis or control selection date per person, number of persons (%) with		
0 values	13,618 (85.7)	75,731 (95.3)
1 value	1801 (11.3)	3257 (4.1)
2+ values	469 (3.0)	447 (0.6)
Closest value in year <u>prior</u> to diagnosis or index date		
Persons with value, n (%)	2270 (14.3)	3704 (4.7)
HbA1c %, mean (SD)	5.9 (0.4)	5.6 (0.4)
Closest value in year <u>after</u> diagnosis or index date		
Persons with value, n (%)	11,990 (75.5)	3839 (4.8)
HbA1c %, mean (SD)	7.5 (2.0)	5.6 (0.4)
LDL cholesterol		
# in year <u>before</u> diagnosis or index date per person, number of persons (%) with		
0 values	10,155 (63.9)	46,485 (58.5)
1 value	4068 (25.6)	23,737 (29.9)
2+ values	1665 (10.5)	9213 (11.6)
Closest value in year <u>prior</u> to diagnosis or index date		
Persons with value, n (%)	5733 (36.1)	32,950 (41.5)
LDL-cholesterol, mg/dL, mean (SD)	107.2 (35.6)	109.6 (33.0)
Closest value in year <u>after</u> diagnosis or index date		
Persons with value, n (%)	11,726 (73.8)	34,223 (43.1)
LDL-cholesterol, mg/dL, mean (SD)	108.5 (36.7)	111.1 (33.7)
Triglycerides		
# in year <u>before</u> diagnosis or index date per person, number of persons (%) with		
0 values	10,529 (66.3)	48,714 (61.3)
1 value	3869 (24.4)	22,585 (28.4)
2+ values	1490 (9.4)	8136 (10.2)
Closest value in year <u>prior</u> to diagnosis or index date		
Persons with value, n (%)	5359 (33.7)	30,721 (38.7)
Triglycerides, mg/dL, mean (SD)	188.7 (131.7)	133.7 (81.2)
Closest value in year <u>after</u> diagnosis or index date		
Persons with value, n (%)	11,207 (70.5)	31,663 (39.9)
Triglycerides, mg/dL, mean (SD)	216.5 (244.8)	135.0 (86.8)
Glucose, fasting		
# in year <u>before</u> diagnosis or index date per person, # of persons (%) with		
0 values	12,139 (76.4)	56,198 (70.8)
1 value	2968 (18.7)	19,023 (24.0)
2+ values	781 (5.0)	4214 (5.3)

5

Supplemental material

BMJ Publishing Group Limited (BMJ) disclaims all liability and responsibility arising from any reliance placed on this supplemental material which has been supplied by the author(s)

BMJ Open

Variable	Cases	Controls
Closest value in year <u>prior</u> to diagnosis or index date		
Persons with value, n (%)	3749 (23.6)	23,237 (29.3)
Glucose, mg/dL, mean (SD)	108.5 (11.8)	95.8 (9.3)
Closest value in year <u>after</u> diagnosis or index date		
Persons with value, n (%)	9259 (58.3)	24,105 (30.3)
Glucose, mg/dL, mean (SD)	147.9 (60.9)	95.9 (9.3)
Glucose, unspecified		
# in year <u>before</u> diagnosis or index date per person, # persons (%) with		
0 values	9913 (62.4)	54,258 (68.3)
1 value	3115 (19.6)	15,293 (19.3)
2+ values	2860 (18.0)	9884 (12.4)
Closest value in year <u>prior</u> to diagnosis or index date		
Persons with value, n (%)	5975 (37.6)	25,177 (31.7)
Glucose, mg/dL, mean (SD)	124.6 (28.2)	97.7 (15.5)
Closest value in year <u>after</u> diagnosis or index date		
Persons with value, n (%)	10,833 (68.2)	27,779 (35.0)
Glucose, mg/dL, mean (SD)	170.7 (95.2)	98.4 (16.5)
Body mass index (BMI)		
# in year <u>before</u> diagnosis or index date per person, mean (SD)	3.1 (4.1)	2.4 (3.2)
Closest value in year <u>prior</u> to diagnosis or index date		
Persons with value, n (%)	11,237 (70.7)	54,733 (68.9)
BMI, kg/m^2, mean (SD)	36.2 (8.4)	29.3 (6.4)
Closest value in year <u>after</u> diagnosis or index date		
Persons with value, n (%)	13,957 (87.9)	65,084 (81.9)
BMI, kg/m^2, mean (SD)	36.0 (8.4)	29.3 (6.4)

Supplemental material

BMJ Publishing Group Limited (BMJ) disclaims all liability and responsibility arising from any reliance
placed on this supplemental material which has been supplied by the author(s)

BMJ Open

Table S5. Adjusted* associations of selected independent variables with type 2 diabetes status stratified by administrative community type.

Variable	Stratified by Administrative Community Type			Stratified by Administrative Community Type		
	Boroughs	City Census Tracts	Townships	Boroughs	City Census Tracts	Townships
	Model 1a OR (95% CI)	Model 1b OR (95% CI)	Model 1c OR (95% CI)	Model 2a OR (95% CI)	Model 2b OR (95% CI)	Model 2c OR (95% CI)
Race						
White	1.0	1.0	1.0	1.0	1.0	1.0
All others	1.44 (1.12, 1.94)	1.30 (1.05, 1.60)	1.36 (1.14, 1.61)	1.43 (1.12, 1.84)	1.28 (1.04, 1.58)	1.35 (1.14, 1.61)
Ethnicity						
Non-Hispanic	1.0	1.0	1.0	1.0	1.0	1.0
Hispanic	1.50 (1.16, 1.94)	1.33 (1.02, 1.72)	1.52 (1.16, 1.97)	1.50 (1.16, 1.94)	1.32 (1.02, 1.71)	1.52 (1.17. 1.97)
Medical Assistance						
< 50% of time	1.0	1.0	1.0	1.0	1.0	1.0
50+% of time	1.66 (1.47, 1.86)	1.46 (1.26, 1.70)	1.83 (1.61, 2.09)	1.66 (1.48, 1.86)	1.48 (1.27, 1.72)	1.83 (1.61, 2.09)
CSD **						
Q1	0.88 (0.77, 1.01)	0.75 (0.56, 1.00)	0.93 (0.84, 1.02)			
Q2	0.96 (0.84, 1.08)	0.77 (0.63, 0.94)	0.97 (0.89, 1.06)			
Q3	0.98 (0.87, 1.10)	0.78 (0.67, 0.91)	0.98 (0.89, 1.07)			
Q4	1.0	1.0	1.0			
NDVI, 1250x1250m †						
T1				1.0	1.0	1.0
T2				0.93 (0.87, 0.99)	0.76 (0.64, 0.90)	0.93 (0.87, 0.99)
T3				0.85 (0.76, 0.96)	0.76 (0.50, 1.17)	0.90 (0.84, 0.96)

Abbreviations: CSD = community socioeconomic deprivation; NDVI = normalized difference vegetation index;

* Logistic regression models using generalized estimating equations with robust standard errors; also adjusted for sex and age (age, age^2, age^3).

** Quartile cutoffs were defined within the three time periods; the range of values for persons in Q1, Q2, Q3, and Q4 were -25.06 to -1.82; -1.99 to 0.10; 0.005 to 2.05; and 1.89 to 12.4, respectively.

† The range of values in T1, T2, and T3 were 0.07 to 0.627, 0.63 to 0.756, and 0.76 to 0.94, respectively.

Supplemental material

BMJ Publishing Group Limited (BMJ) disclaims all liability and responsibility arising from any reliance placed on this supplemental material which has been supplied by the author(s)

BMJ Open

Table S6. Adjusted* associations of selected independent variables with type 2 diabetes status stratified by administrative community type with county and community socioeconomic deprivation **_OR_** greenness.

	Stratified by Administrative Community Type		
	Boroughs	City Census Tracts	Townships
Variable	Model 1 OR (95% CI)	Model 1 OR (95% CI)	Model 1 OR (95% CI)
Model 1 – with county and community socioeconomic deprivation (CSD)			
Race			
White	1.0	1.0	1.0
All others	1.45 (1.13, 1.86)	1.31 (1.06, 1.62)	1.39 (1.16, 1.66)
Ethnicity			
Non-Hispanic	1.0	1.0	1.0
Hispanic	1.49 (1.15, 1.92)	1.32 (1.02, 1.71)	1.55 (1.18, 2.04)
Medical Assistance			
< 50% of time	1.0	1.0	1.0
50+% of time	1.66 (1.47, 1.87)	1.48 (1.28, 1.72)	1.85 (1.62, 2.11)
Community socioeconomic deprivation, quartiles			
Q1	0.87 (0.76, 0.996)	0.71 (0.52, 0.95)	0.91 (0.82, 0.99)
Q2	0.93 (0.83, 1.06)	0.78 (0.65, 0.95)	0.96 (0.88, 1.05)
Q3	0.97 (0.87, 1.09)	0.79 (0.67, 0.93)	0.98 (0.90, 1.07)
Q4	1.0	1.0	1.0
County			
Luzerne	1.0	1.0	1.0
Blair	**0.64** (0.51, 0.81)	0.62 (0.23, 1.64)	0.86 (0.61, 1.21)
Clearfield	1.00 (0.82, 1.24)	**0.76** (0.66, 0.87)	0.97 (0.82, 1.15)
Dauphin	0.90 (0.56, 1.45)	**2.81** (1.47, 5.37)	1.43 (0.96, 2.15)
Juniata	**1.68** (1.22, 2.31)	NA†	1.18 (0.99, 1.41)
Lackawanna	1.12 (0.96, 1.37)	**1.23** (1.06, 1.43)	1.13 (0.93, 1.38)
Lehigh	**18.2** (2.00, 165.1)	2.00 (0.85, 4.68)	0.66 (0.26, 1.65)
Mifflin	**1.20** (1.00, 1.43)	NA	1.06 (0.93, 1.21)
Monroe	**0.73** (0.59, 0.91)	NA	**0.85** (0.74, 0.98)
Perry	**3.16** (1.34, 7.47)	NA	0.96 (0.51, 1.83)
Potter	**4.90** (4.42, 5.43)	NA	0.71 (0.15, 3.31)
Schuylkill	0.91 (0.80, 1.02)	0.93 (0.80, 1.07)	**0.82** (0.73, 0.91)
Snyder	**0.84** (0.72, 0.98)	NA	1.01 (0.88, 1.16)
Sullivan	0.63 (0.38, 1.07)	NA	**0.65** (0.47, 0.90)
Union	0.84 (0.53, 1.34)	NA	**0.80** (0.66, 0.98)
Wayne	**3.36** (1.83, 6.16)	NA	0.96 (0.59, 1.58)
Wyoming	**0.86** (0.76, 0.96)	NA	**1.15** (1.00, 1.32)
Model 2 – same as Model 1, but with NDVI not CSD, with county; only NDVI associations are shown			
Normalized difference vegetation index (NDVI)			
T1	1.0	1.0	1.0
T2	0.91 (0.85, 0.98)	0.77 (0.64, 0.92)	0.93 (0.87, 0.99)
T3	0.85 (0.75, 0.97)	0.76 (0.48, 1.19)	0.90 (0.84, 0.97)

* Logistic regression models using generalized estimating equations with robust standard errors; also adjusted for sex and age (age, age^2, age^3). Counties with at least one association that excluded 1.0 in confidence interval included in table (37 counties were included in total; 36 county indicators vs. Luzerne County as reference).

† NA = these counties did not have city minor civil divisions or did not converge due to small numbers.