

Supplemental Table 3. Sensitivity Analysis using Univariate Regression

Individual DM Status – Univariate Logistic Regression Results				
<i>Category</i>	<i>Factor</i>	<i>Mean (SD) [Range]</i>	<i>Coefficient (SE)</i>	<i>AUROC</i>
Demographic	Age	49.4 (15.9) [20.0-80.0]	0.0441 (0.00231)*	0.675
	Female	50.9	-0.156 (0.0836)*	0.589
	African American	10.2	0.512 (0.0820)*	0.612
	Hispanic	14.2	0.070 (0.080)	0.507
Socioeconomic	Household Income	NA	-0.935 (0.163)*	0.593
	Some College	61.1	-0.421 (0.0830)*	0.610
County-Level DM Prevalence – Univariate Linear Regression Results				
<i>Category</i>	<i>Factor</i>	<i>Mean (SD) [Range], %</i>	<i>Coefficient (SE)</i>	<i>R²</i>
Demographic	Percent < 18 years old	22.3 (3.5) [0.0-40.9]	0.0452 (0.0129)	0.00389
	Percent > 65 years old	18.4 (4.6) [4.6-56.3]	0.128 (0.00957)*	0.0543
	Percent Female	49.9 (2.3) [27.8-56.6]	0.226 (0.0194)*	0.0417
	Percent African American	9.0 (14.3) [0.0-85.2]	0.0792 (0.00280)*	0.203
	Percent Hispanic	9.3 (13.7) [0.5-96.3]	-0.0521 (0.00316)*	0.0799
Socioeconomic	Median Household Income	91.3 (2.1) [84.6-100.0]	-0.739 (0.0174)*	0.367
	Percent Some College	57.2 (11.6) [15.1-94.1]	-0.115 (0.00331)*	0.276

Univariate regression results for models using factors shared between NHANES (individual-level) and CHR (county-level) datasets applying the same definition of DM (i.e. self-reported DM only) for both datasets. For female gender, Hispanic and African-American race/ethnicity, and education level factors from the individual-level (NHANES) data, summary characteristics are expressed in terms of percent of total sample; not all summary statistics could be calculated. Similarly, for household income from the individual-level data, values were collected and stored as ranges of income; summary statistics could not be calculated. Finally, for county-level median household income, variables were normalized and scaled to have a maximum value of 100.

* $p < 0.001$

† $p < 0.05$