Supplemental Online Content

eMethods

This supplemental material has been provided by the authors to give readers additional information about their work.

eMethods

Data sources: Patient data came from the Komodo Healthcare Map – a nationally representative, longitudinal, and de-identified US pharmacy and medical claims database. Patient encounter data for 207 million individuals in the Komodo Healthcare Map were derived directly from payer sources, including 100% fully integrated fee-for-service Medicare data, Medicare Advantage, commercial, and Medicaid claims. The dataset is representative of the US insured population from all 50 US states and provides comprehensive pharmaceutical, socio-demographic, diagnostic, and resource utilization information.

GLP-1 discontinuation: GLP-1 discontinuation was assessed at 3, 6, and 12 months after the index date, defined as not having any GLP-1 fill in the following 135 days. Discontinuation at 3 months indicates that for 135 days *after* 3 months from the index prescription, there were no further prescriptions for a GLP-1 agonist. Similarly, discontinuation at 6 and 12 months indicated that there was no GLP-1 after 6 months, and 12 months respectively. We did not require continuous prescription for 3 months, 6 months, or 12 months. The 3-month discontinuation simply indicated that at 3 months, the patient got no additional GLP-1 in the following 135 days. Thus, a patient that filled only one prescription (the index prescription) and did not subsequently fill another GLP-1 would fall into the 3-month discontinuation group. Note that this patient may also fall into the discontinuation-by-12-months group since they would not have any GLP-1 fill in the 135-day time period after 12 months.

Baseline conditions:

Conditions	Definition
Type 2 diabetes	Patients were classified as having type 2 diabetes (T2D) if they had 2+ encounters with any diagnosis code for T2D (ICD-10 codes E11*), 2+ pharmacy claims for any antidiabetic medication other than GLP-1, or any A1C result ≥ 6.5% at baseline.
Obesity	ICD-10 codes: E66.0*, E66.1, E66.2, E66.8, E66.9, Z68.3*, Z68.4*
Heart failure	ICD-10 codes: I09.81, I11.0*, I13.0*, I13.2*, I50*, I5A*
Cardiovascular disease (except heart failure)	ICD-10 codes: I20*, I21*, I22*, I23*, I24*, I25*, I26*, I42*, I43*, I44*, I60*, I61*, I62*, I63*, I64*, I65*, I66*, I67*, I68*, I69*, I80*, I81*, I82*
Chronic kidney disease	ICD-10 codes: I12*, I13*, N18*
Gastrointestinal (GI) diseases	ICD-10 codes: K52.9, K56*, K59.0*, K59.1, K59.2, K59.3*, K85.0*, K85.3*, K85.8*, K85.9*, K80*, K81*, K86.1, K86.8*, K86.9, K87*, R10*, R11*, R12*, R15*, R19.4, R19.5, R19.7

Social determinants of health (SDoH) index:

Socioeconomic conditions in a community describe important aspects of social organization, structure, stratification, and/or environment where individuals live. These aspects together mold and influence the health aspects of a person living in those communities. One way to characterize a community is to use census indicators representing the multidimensional aspect of a community and combine them to arrive at a composite score. The SDOH index in this study is a composite metric at the census tract level derived from 6 different domains. A lower score represents communities with a lower level of social determinants than a higher scored community. The overall SDOH score was categorized into four

mutually exclusive groups based on its quantile distribution – including low, medium, high, and very high social needs.

The scoring logic was based on 22 different derived metrics (listed in the table below). In short, clusters were created in a way that variables from the same cluster were highly correlated with each other but had a low correlation with any other cluster. Once cluster labels were assigned to observations, average values for each input variable were calculated per each segment. Finally, we combined each average input value per segment together and created a variable to be used in weight attribution for the input variables. Ordinary least squares (OLS) regression was used to extract the contribution weights from each of the input variables. The slope coefficients were proxies for weights attributed to each input and after converted to standardized estimates, they became the final weights.

Domains	Characteristics	Data sources
Economics	Proportion of households below poverty line	2014-2018 ACS
	Proportion of total households with zero income	
	Number of businesses per capita	
	Median household income relative to county median household income	
	Unemployment rate	
Education	Proportion of total population with no high school diploma	2014-2018 ACS
	Proportion of total population that dropped out of college	
	Health literacy ratio relative to the national average	
Food access	Share of population beyond 1 mile (urban) or 10 miles (rural) from a supermarket	2014-2018 ACS
	Share of low income individuals that are beyond 1 mile (urban) or 10 miles (rural) from a supermarket	
	Share of households without a vehicle that are beyond 1 mile (urban) or 10 miles (rural) from a supermarket	
	Share of households receiving benefits from the Supplemental Nutrition Assistance Program (SNAP) that are beyond 1 mile (urban) or 10 miles (rural) from a supermarket	
	Walkability score	EPA
Health	Proportion of non-insured population ages 0-17	2014-2018 ACS
	Proportion of non-insured population ages 18-34	
	Proportion of non-insured population ages 35-64	
	Proportion of non-insured population ages >64	
Infrastructure	Proportion of total population who are homeowners	2014-2018 ACS
	Proportion of total housing units that are vacant	
	Proportion of total population using sustainable means for transportation	

Domains	Characteristics	Data sources
	Ratio of median house value relative to national median sale price year 2014-2018	
Language and culture	Proportion of non-English speakers in total population	2014-2018 ACS

Notes: ACS, American community survey; EPA, US Environmental Protection Agency.