The Hilltop Institute
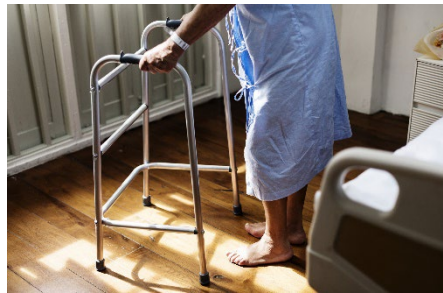## UMBC

Maryland Primary Care Program (MDPCP) Pre-AH Risk Score Specifications and Codebook

October 3, 2019

Version 1

Maryland
DEPARTMENT OF HEALTH

Acknowledgements

Maryland Primary Care Program (MDPCP)
Pre-AH Risk Score Specifications and Codebook

Table of Contents

Maryland
DEPARTMENT OF HEALTH

The Hilltop Institute

1. Overview

The Maryland Primary Care Program (MDPCP) is a key element of the Total Cost of Care (TCOC) All-Payer Model, an agreement between the Centers for Medicare & Medicaid Services (CMS) and the state of Maryland. The MDPCP is a voluntary program that provides funding and support for the delivery of advanced primary care throughout the state. It allows primary care providers to play an increased role in the prevention and management of chronic disease, as well as in the prevention of unnecessary hospital utilization.

As an important part of supporting providers in their care management efforts, the MDPCP will provide to participating practices risk scores of their attributed beneficiaries according to each patient's risk of incurring a potentially avoidable hospitalization or emergency department (ED) visit. Accordingly, The Hilltop Institute has developed the Hilltop Pre-AH (Predicting Avoidable Hospitalizations) Model™ in order to operationalize these risk scores. These patient-level risk scores will be vended to participating medical practices on a monthly basis starting in October 2019 via the MDPCP portal on the Chesapeake Regional Information System for our Patients (CRISP) unified landing page.

This document is intended to explain the intended use, technical implementation, and model performance of the Hilltop Pre-AH Model™ as of August 2019. It will be updated as future versions of the model become operational.

2. Intended Use

The Hilltop Pre-AH Model™ risk scores are intended to add value to the primary care transformation process by facilitating improved efficiency in the allocation of scarce care coordination resources. Theoretically, if such resources are limited and the patients in a given practice panel differ in the benefit they would obtain through care coordination, then patient outcomes are optimized by focusing those care coordination resources on the patients for whom these resources will generate the most benefit.[1] Hilltop's model will be used to rank attributed beneficiaries in each practice's panel based on their risk of experiencing an avoidable hospital event in order to assist in the identification and care coordination efforts for those high-risk individuals.

Hilltop conceptualizes benefit, in this context, as the avoidance of a patient-specific adverse event. Many distinct adverse events are possible (ranging from disease onset to institutionalization to death), but given the emphasis of the MDPCP on the reduction of

---

[1] There is some evidence to suggest that different patients receive different benefits from care coordination services. Researchers have found that proactive care coordination interventions for patients with a high risk of hospitalization have so far led to reductions in avoidable hospitalizations, ED utilization, and readmissions for the Medicaid population but not the Medicare population (Berkowitz et al., 2018).

unneeded utilization, the risk model focuses on *potentially avoidable* hospitalization or ED visits.[2] While this is a composite measure of eleven distinct underlying disease states,[3] each of which entails distinct patient-specific costs, Hilltop treats these events as homogeneous and therefore focuses on patients' *probabilities* of incurring avoidable hospital events. This forms the theoretical foundation for the Hilltop Pre-AH Model™: those individuals with the highest *probability* of incurring an avoidable hospitalization or ED visit are likely to benefit the most from advanced primary care services with respect to that outcome. Through the dissemination of the Hilltop Pre-AH Model™ risk scores, Hilltop aims to facilitate the identification of these individuals within each practice so that practices can allocate their care management resources accordingly.

It is crucial that the risk scores are as accurate as possible: ideally, the riskiest individuals as identified by the model have the highest actual likelihood of incurring an avoidable hospitalization in the next month, and the individuals identified by the model as lowest risk have the lowest actual likelihood of incurring an avoidable hospitalization in the next month. Due to the nature of the modeling problem—estimating the distribution of risk, rather than binary classification—it is not appropriate to use the traditional Receiver Operator Characteristic curve as a measure of model fit.[4] Instead, the utility of the model is assessed using *concentration curves,* which estimate the share of all avoidable hospital events occurring within the riskiest patients. Concentration curves can indicate, for example, that 50 percent of all patients who experience an avoidable hospital event are in the top 10 percent riskiest patients as estimated by the Hilltop Pre-AH Model™. Concentration curves and month-by-month summary scores for the model are presented in Section 4.2, below.

2.1 Differentiation from CMS HCC Risk Scores

It is important to note that the Hilltop Pre-AH Model™ risk scores are conceptually distinct from the CMS Hierarchical Condition Category (HCC) risk scores that are currently presented in CRISP. The Hilltop Pre-AH Model™ risk scores use risk factors based on diagnoses, procedures, medications, utilization, demographics, and geographic factors in order to produce a practice-specific ranking of patient risk of avoidable hospital events within the next month. The CMS HCC risk scores are based on a model that uses diagnosis codes and a limited set of demographic information from a base year in order to predict *expenditures* over the following year. There is likely to be some overlap among individuals who incur an avoidable hospitalization and individuals who experience high medical spending, but the overlap is unlikely to be complete.[5]

---

[2] Potentially avoidable hospitalizations/ED visits are those incurred for medical conditions or diagnoses "for which timely and effective outpatient care can help to reduce the risks of hospitalization by either preventing the onset of an illness or condition, controlling an acute episodic illness or condition, or managing a chronic disease or condition" (Billings et al., 1993). This measure is discussed in greater detail in Section 3.2.1.

[3] See Section 3.3.1 for further information.

[4] For additional detail on this point, see Section 4.2.

[5] Internal testing shows a limited degree of substitutability between the two sets of risk scores. Specifically, we find that the Hilltop Pre-AH Model™ outperforms the CMS HCC risk score in predicting avoidable hospitalization in the following month: of the top 10 percent riskiest individuals ranked by each risk score, the Hilltop Pre-AH Model™ correctly identifies 45-50 percent of all avoidable hospital events, while the CMS HCC risk score identifies approximately 30 percent. Both concentration curves are presented in Section 4.2, below.

The Hilltop Institute

High medical expenditure can reflect a number of factors ranging from moderate utilization of high-cost procedures, high utilization of moderate-cost procedures, underlying morbidity, or geographic differences in treatment or referral practices.

Moreover, the theoretical interpretation of each risk score differs substantially. The CMS HCC risk score was developed as a capitated payment risk adjustment methodology for Medicare Advantage participants in order to "address [the] issue of risk selection and to compensate Medicare Advantage health plans for accepting the risk of enrolling beneficiaries of varying health statuses" (CMS, 2018, pp. 9-10). Additionally, "the underlying risk assessment is designed to accurately explain the variation at the group level, not at the individual level, because risk adjustment is applied to large groups" (CMS, 2018, pp. 9-10). Note that "risk" for the CMS HCC risk model refers to *actuarial* risk: this model seeks to predict average expenditures over large groups of individuals. In contrast, the Hilltop Pre-AH Model™ risk score is designed to estimate, as closely as possible, event risk: that is, an *individual's* risk of an avoidable hospital event in the following month.

There are also differences in the time horizons of each risk score. CMS HCC "final risk scores are generally available 16-18 months after the close of the base year. For example, 2017 risk scores (based on 2016 diagnoses) will be available in the spring of 2018" (Center for Medicare & Medicaid Innovation, 2017, p. 26). The Hilltop Pre-AH Model™ risk scores, however, are updated monthly and use patient-level risk factor information current to the most recently available month of Medicare claims in order to generate risk scores. This is a strength of the Hilltop Pre-AH Model™: these risk scores reflect the underlying patient condition with a lag of only, at most, two months.[6] Finally, by definition, avoidable hospital events are preventable through timely primary care and so, in principle, the identification and management of individuals at high risk of incurring an avoidable hospitalization may result in the avoidance of that particular hospitalization event. High medical expenditures, however, may reflect underlying morbidities that would necessitate utilization *regardless* of primary care intervention.

2.2 Clinical Vignette

In order to illustrate the intended use of the Hilltop Pre-AH Model™ risk scores, we have created a hypothetical clinical vignette for an MDPCP Track 1 practice. For the sake of exposition, the patient panel consists of thirteen patients, each of which represents ten similar patients. Table 1 displays the patient panel, along with each patient's (hypothetical) Hilltop Pre-AH Model™ risk score and CMS Risk Tier.

---

[6] This lag depends on the timing of two factors: the receipt of CCLF claims by Hilltop, and the publication of the risk scores by CRISP. For example, claims data delivered to Hilltop in mid-August 2019 reflect utilization through mid-July 2019. Hilltop uses these July 2019 risk factors to generate risk scores that reflect the likelihood of incurring an avoidable hospital event in August 2019. These scores, then, will be vended to participating practices in the CRISP update in mid-September 2019.

Table 1. Hypothetical Patient Panel

Patient Name	Pre-AH Risk Score (%)	CMS Risk Tier
Patient 1	75%	Complex[7]
Patient 2	15%	Complex
Patient 3	5%	Tier 4
Patient 4	4%	Complex
Patient 5	2%	Tier 3
Patient 6	1%	Tier 3
Patient 7	Less than 1%	Tier 2
Patient 8	Less than 1%	Tier 2
Patient 9	Less than 1%	Tier 1
Patient 10	Less than 1%	Tier 2
Patient 11	Less than 1%	Tier 1
Patient 12	Less than 1%	Tier 1
Patient 13	Less than 1%	Tier 1

Patients in this practice are listed in descending order of risk. Based on the most recently available month of risk factors spanning diagnoses, procedures, medications, utilization, demographics, and geographic information, in conjunction with risk coefficients derived from training data, Patient 1 (or, equivalently, the ten patients represented by Patient 1) has a 75 percent chance of incurring an avoidable hospital event next month.[8] Patient 2 is the next riskiest, and has a 15 percent chance of incurring an avoidable hospital event. Patient 3 is the next riskiest, with a 5 percent chance. The distribution of risk is highly skewed: the majority of the practice's panel has less than 1 percent chance of incurring an avoidable hospitalization in the following month, and all but two of the patients have under a 6 percent event risk.[9]

Based on the MDPCP Care Management Fee (CMF) structure, this practice would receive $2,600 each month.[10] Distributing the CMF revenue equally across all 130 underlying patients would

---

[7] It is important to note that while the CMS risk tier is correlated with Hilltop Pre-AH Model™ risk scores, the correlation is not perfect for two reasons: first, CMS risk tiers are based on underlying HCC score, which is conceptually distinct from the Pre-AH risk score. Second, certain groups of patients are automatically assigned to certain CMS risk tiers, which further reduces the correlation between the two measures. In particular, beneficiaries without sufficiently long clinical histories are assigned to CMS risk tier 2, while beneficiaries with "a diagnosis of dementia, substance use disorder, or severe and persistent mental illness" are assigned to the Complex tier, regardless of their HCC score (CMMI 2019). These individuals may, in turn, have relatively low (or high) risk of avoidable hospitalizations, meaning that an individual in, for example, the Complex CMS risk tier may have a low Pre-AH risk score. We highlight this point in Table 1 by presenting a non-monotonic relationship between Pre-AH risk score and CMS risk tier.

[8] See Section 3.2 for a more detailed discussion of the training and scoring process.

[9] While the data for this clinical vignette are hypothetical, the Hilltop Pre-AH Model™ risk scores are, in actuality, even more skewed: the average probability of incurring a future hospitalization is roughly 0.5 percent, while the maximum probability in the MDPCP cohort is greater than 99 percent.

[10] $50 for each of the 30 patients in the Complex tier; $30 for each of the 10 patients in Tier 4; $16 for each of the 20 patients in Tier 3; $8 for each of the 30 patients in Tier 2; and $6 for each of the 40 patients in Tier 1. For the purposes of this clinical vignette, we do not account for the Performance-Based Incentive Payment (PBIP), although this would potentially add $325 per month to this practice's MDPCP revenues.

The Hilltop Institute

result in each patient receiving $20.00 of advanced primary care services each month. This distribution is unlikely to have a significant impact on patient outcomes: the low-risk individuals would be low-risk even without the advanced primary care intervention, and the high-risk individuals may require more resource-intensive interventions in order to experience improvement in outcomes.[11]

2.3 Business Process

Hilltop remains agnostic as to the particular types of interventions that are best-suited for the high-risk MDPCP population. Many interventions are possible, ranging from medication reconciliation to patient education to scheduling assistance, and patients are likely to respond best to different interventions based on their clinical and social needs. Interested readers should see published best practices in care coordination and care management.[12] Whatever the intervention strategy, Hilltop recommends that care managers and other users of the Hilltop Pre-AH Model™ risk score allocate their effort first to individuals with the highest risk of incurring an avoidable hospital event in the following month. This risk score is not, however, meant to override the clinical and subject matter expertise of the practice or their care transformation organization (CTO) partners and should be used in conjunction with the practice's current care coordination protocols. For details on the user interface of the Hilltop Pre-AH Model™ risk scores, readers should see the CRISP user manual.[13]

2.4 Reason for Risk

Currently, the Hilltop Pre-AH Model™ risk scores are individuals' probabilities of incurring an avoidable hospitalization in the following month. This probability is derived from a rich data set of risk factors—ranging from demographics to clinical diagnoses to geographical and social determinants of health—to which risk coefficients from a training model are applied. Each attributed MDPCP beneficiary will receive an updated score each month that reflects new risk factor information derived from Medicare claims in the previous month.

In a release planned for later this year, the Hilltop Pre-AH Model™ risk score will be accompanied by the three largest "reasons for risk" for each patient in a practice. The practice-specific risk score view on CRISP will be augmented to include those risk factors which contribute the most to each patient's risk of incurring an avoidable hospital event in the following month. This information is intended to allow for the implementation of patient-specific, near-to-real-time,

---

[11] Liaw et al. (2015) conclude that, based on a review of four CMS-funded demonstrations involving care management fees, "to generate savings, resource allocation cannot be homogeneous. Instead, practices must focus more intensely on those at highest risk of utilization" (p. 557). Indeed, this may (partly) explain the varying effectiveness of care management, care coordination, and intensive primary care interventions as documented in the academic literature; many patients have low underlying risk of adverse outcomes, thus obviating the need for intervention, and the few high-risk patients may require significant intervention resources. For summaries of the literature on this subject, see Edwards et al. (2017) and Baker et al. (2018).

[12] See examples at Hong et al. (2014); McCarthy et al. (2015); and Anderson et al. (2015).

[13] https://crisphealth.org/resources/training-materials/

tailored interventions based on statistically validated risk factors. This functionality is currently in development.

3. Technical Implementation

This section presents details on data sources, risk factors, and methodology.

3.1 Data Sources

The Hilltop Pre-AH Model™ relies largely on data from Claim and Claim Line Feed (CCLF) Medicare claims files, supplemented with various publicly available environmental data sets used to generate the environmental risk factors. These data sources are detailed below.

3.1.1 CCLF Data

The majority of the risk factors in the Hilltop Pre-AH Model™ are derived from CCLF Medicare Parts A, B, and D claims files. Each month, Hilltop receives Part A claims, Part A revenue centers, Part A procedure codes, Part A diagnosis codes, Part B claim lines, Part B durable medical equipment claims, Part D claims, and patient demographic information (which also includes eligibility information) from CMS.[14] Additionally, Hilltop receives beneficiary attribution files and practice rosters each quarter.

Upon receipt of the monthly claims files, Hilltop first performs automated data validity checks in order to assess the integrity of the CCLF data files, followed by a data reduction step that subsets the claims files against the beneficiary attribution file. The resulting files retain the raw claims data that are inputs to the risk factor feature engineering process, but discard the claims for individuals that are not in the MDPCP population. The resulting data comprises approximately 220,000 individuals across almost 400 practices. These individuals incurred approximately 2.1 million part A claims, 16.8 million part B claims, and 12.4 million part D claims in the three-year period of February 2016 to January 2019.

Using SAS 9.4, Hilltop creates the model using risk factors identified in the literature review.[15] The risk factors are described in Section 3.2 and in greater detail in Appendix 1.

3.1.2 Social Determinants of Health Data Set

In order to control for environmental factors that may affect patients' probabilities of incurring avoidable hospitalizations, the risk model includes a rich set of area-level covariates derived from publicly available sources. Based on the "beneficiary ZIP code as per Medicare enrollment"

---

[14] For detailed documentation of these files, please see "Maryland Primary Care Program (MDPCP) CRISP Extract" (June 2019).

[15] Certain risk factors identified in the literature review were not ultimately operationalizable in Phase 1 of the Hilltop Pre-AH Model™. We will incorporate additional risk factors in future iterations of the model.

(BENE_ZIP_CD), each attributed beneficiary is linked to environmental characteristics in his or her residential area.

It is important to note that ZIP codes, which are generated by the United States Postal Service, do not represent polygonal shapes; instead, they represent collections of mailing addresses. The U.S. Census Bureau uses an analogous concept—the ZIP Code Tabulation Area (ZCTA)—in order to publish data tables at a ZIP code-like level of granularity.

ZCTAs are approximate area representations of U.S. Postal Service 5-digit ZIP code service areas that the Census Bureau creates using census blocks to present statistical data from censuses and surveys.[16] The Census Bureau defines ZCTAs by allocating each block that contains addresses to a ZCTA, usually to the ZCTA that reflects the most frequently occurring ZIP Code for the addresses within that block. Blocks that do not contain addresses but that are completely surrounded by a single ZIP code tabulation area (enclaves) are assigned to the surrounding ZCTA; those surrounded by multiple ZCTA will be added to a single ZCTA based on the longest shared border. ZIP Codes that cover primarily nonresidential or P.O. box addresses may not have a corresponding ZCTA because the delineation process uses primarily residential addresses. The Census Bureau (2018a; 2018b) notes that "in most instances the ZCTA code is the same as the ZIP code for an area."

Of the 33,120 ZCTAs in the United States, 24.04 percent are missing data for at least one variable of the Hilltop Pre-AH Model™ ZIP code-level data set.[17, 18] This issue is ameliorated, although still present, in the estimation sample of Medicare beneficiaries attributed to MDPCP-participating practices: of the 219,173 individuals in the baseline beneficiary demographic file with a valid ZIP code, at least one variable in the ZIP code-level data set is missing information for 0.75 percent of individuals.[19]

In order to address this issue, Hilltop leverages the informational content of ZIP codes to impute the missing values of these ZIP code-level variables. Within each five-digit ZIP code, the first three digits represent the "Sectional Center Facility," a centralized mail distribution hub that sorts and distributes mail to local post offices according to the fourth and fifth digits in the ZIP code. (Congressional Research Services, 2006; U.S. Postal Service Office of Inspector General, 2013). For example, consider ZIP code 21250, which is located in Catonsville, MD. The first three digits indicate that this is served by Sectional Center Facility 212. All other ZIP codes beginning with 212, then, are *also* served by that particular Sectional Center Facility, implying that they are

---

[16] Note: USPS ZIP codes are not areal features but a collection of mail delivery routes.

[17] It is important to note that two variables—"Social Workers per 1,000 population" and "Percent Physician Diversity"—are only available for a subset of counties due to geographical identification in underlying American Community Survey data. In the data creation process, we impute the missing values for these variables as using the geographically *unidentified* value for each variable within a given state. This is discussed in greater detail below.

[18] While ZIP codes and ZCTAs capture different underlying geographic concepts, they are used interchangeably for the purposes of this discussion.

[19] The *pp23* beneficiary demographic file contains 219,522 distinct individuals. Of these, 349 do not have a ZIP code that merges on to the ZCTA-level data set (usually due to beneficiaries using P.O. boxes as their address of record, which are not attributed to ZCTAs). The ZCTA is imputed for these 349 as described in Section 3.3.2.

The Hilltop Institute

relatively spatially proximal to ZIP Code 21250. These spatially proximal ZIP codes are the basis for the imputation procedure. Specifically, Hilltop imputes missing data for a given variable in a given ZIP code as the population-weighted average value of that variable for all non-missing ZIP codes within a given Sectional Center Facility.

To fix ideas, consider three ZIP codes within a given Sectional Center Facility code: 55501, 55502, and 55503. Suppose that the value of the variable "Percent of Population Aged 65+" is missing for ZIP code 55503 but is 20 percent for 55501 and 10 percent for 55502. Additionally, suppose that ZIP code 55501 contains 10,000 individuals, and ZIP code 55502 contains 1,000 individuals. The imputed value for ZIP code 55503 is (10,000/11,000)*20 percent + (1,000/11,000)*10 percent = 19.1 percent. This imputed value is almost certain to contain noise: the weighted average will not exactly equal the unobserved value of this variable for ZIP code 55503. However, to the extent that spatially proximal ZIP codes (that is, within a given Sectional Center Facility catchment area) are *similar* in terms of observable characteristics, this imputation method balances computational feasibility with accuracy.[20]

Twenty-three ZIP codes are located in Sectional Center Facility catchment areas (202, 204, 205, 753, and 772) with zero total population.[21] These ZIP codes appear to be attributed to specific buildings in urban areas (for example, the Federal Bureau of Investigation building, in 20535). The imputation procedure fails for these ZIP codes because it is impossible to calculate population-weighted averages for the ZIP codes within those areas. Consequently, all missing values for these particular ZIP codes are imputed as 0.

Appendix 1 lists the general data sources for each of the risk factors. See below for details of the data sources for these environmental risk factors.

- ACS data are from the American Community Survey. ZCTA data were obtained through Census Bureau's American FactFinder (https://factfinder.census.gov/faces/nav/jsf/pages/index.xhtml). Data are at granularity of ZCTA.

- IRS SOI data are from the Internal Revenue Service Statistics of Income. ZCTA data were obtained through https://www.irs.gov/statistics/soi-tax-stats-individual-income-tax-statistics-2016-ZIP-code-data-soi. Data are at granularity of ZIP Code.

- USDA rural-urban commuting data are Version 3.10 of the ZIP code Rural-Urban Commuting Areas (RUCA) taxonomy. These data are comprised of 10 codes that "delineate metropolitan, micropolitan, small town, and rural commuting areas based on the size and direction of the primary (largest) commuting flows." Census tract level data and documentation are here: https://www.ers.usda.gov/data-products/rural-urban-commuting-area-codes/. Researchers at the University of North Dakota have published a

---

[20] In future phases of MDPCP, Hilltop will seek to improve the accuracy of imputation through the use of geographic software and more sophisticated spatial imputation techniques.

[21] The ZIP Codes are: 20202, 20204, 20228, 20230, 20240, 20245, 20260, 20405, 20418, 20427, 20506, 20510, 20520, 20535, 20540, 20551, 20553, 20560, 20565, 20566, 20593, 75390, and 77201.

ZIP code-level data set here: https://ruralhealth.und.edu/ruca. The ZIP code-level data are used in the Hilltop Pre-AH Model™.

- Neighborhood Atlas data are from the University of Wisconsin School of Medicine. 2015 Area Deprivation Index (ADI) data were obtained at the Block Group level from https://www.neighborhoodatlas.medicine.wisc.edu/download. The ADI is a national ranking (from 1 to 100) of census block groups by socioeconomic disadvantage. Lower scores indicate less disadvantage, and higher scores indicate more disadvantage. Hilltop used a ZCTA block group crosswalk obtained from http://mcdc.missouri.edu/applications/geocorr2014.html in order to link the block group level deprivation index (from Neighborhood Atlas) to the ZCTA level. ZCTA-level deprivation scores were estimated as weighted averages of block group deprivation scores, with weights being the block group's share of the ZCTA's 2010 population.

- CMS provider locations are from the December 2018 Public Use Provider of Services file (https://www.cms.gov/Research-Statistics-Data-and-Systems/Downloadable-Public-Use-Files/Provider-of-Services/). The ZIP codes for active providers were extracted and merged on to the Census ZCTA template. Hilltop used active short-term/critical access/transplant hospitals for its hospital-based risk factors.

- Veterans Affairs provider locations are from the VA directory (https://www.va.gov/directory/guide/rpt_fac_list.cfm). Locations are retained if the name of the facility contains the term "Clinic" or "Medical Center." ZIP5 was extracted from the address field.

- AMA data are "Census tract layer attributes for American Medical Association Primary Care Physician Data, 2011," published by the Health Resources and Service Administration (HRSA) data warehouse. Specific source: https://data.hrsa.gov/DataDownload/PCSA/2010/t_ama2011_060614.dbf. Data are at the granularity of Census tract. Hilltop used a ZCTA-Census tract crosswalk obtained from http://mcdc.missouri.edu/applications/geocorr2014.html in order to link to the tract-level number of primary care and specialist physicians to the ZCTA level. Hilltop calculated the ZCTA-level counts of primary care and specialty physicians as weighted averages of census tract counts, with weights being the share of the tract's population that is in each ZCTA.

- Land area is from the 2018 Census Gazetteer (https://www.census.gov/geographies/reference-files/time-series/geo/gazetteer-files.html). Area is in square miles.

- Area Health Resources File (https://data.hrsa.gov/data/download) contains county-level data on a variety of health-related topics. Hilltop links this to ZCTAs using a ZCTA-county crosswalk (available from https://www2.census.gov/geo/docs/maps-data/data/rel/zcta_county_rel_10.txt). Counties are mapped to ZCTAs by attributing to each ZCTA the features of the corresponding county within which the ZCTA lies. In the cases where a ZCTA is in two or more counties, ZCTA-level characteristics are estimated using a weighted average of county attributes, weighted by the fraction of the population

of the ZCTA that lies in each county. For example, if 80 percent of people in ZCTA #1 live in County #1, and 20 percent of people in ZCTA #1 live in County #2, and if County #1 is a primary care shortage area, and County #2 is not a primary care shortage area, then ZCTA #1 will receive a value of .8 * 1 + .2*0 = .8 for this variable. The intuition is that, assuming that population is uniformly distributed across the ZCTA, it is reasonable to expect 80 percent of the residents of that ZCTA to live in a primary care shortage area.

- ACS individual-level data are from IPUMS (https://usa.ipums.org/usa/index.shtml). Individual-level microdata are filtered to retain only certain occupations and then aggregated to the county level. This is disaggregated from the county level to the ZCTA level by attributing to each ZCTA the features of the corresponding county within which the ZCTA lies. In the cases where a ZCTA is in two or more counties, ZCTA-level characteristics are estimated using a weighted average of county attributes, weighted by the fraction of the population of a particular ZCTA that lies in each county.

   The variables derived from this data set—"Social Workers per 1,000 population" and "Percent Physician Diversity"—are populated only for a subset of counties covering approximately 1/3 of ZCTAs nationally. This is due to county identification in the underlying source data: the American Community Survey microdata does not publically identify counties for respondents. Instead, IPUMS—an organization based at the University of Minnesota that cleans, documents, and integrates data across publicly available data sets—identifies counties, where possible, from other low-level geographic identifiers and all remaining unidentified counties within a state are aggregated together (Ruggles et al., 2019).[22] For example, of the 24 counties in Maryland (including Baltimore City), only 12 counties have county-specific values for "Percent Physician Diversity." The value for the remaining 12 counties is 10.22 percent.

   Hilltop imputes the missing county-level values for these two variables using the geographically *unidentified* value for each variable within a given state. To continue the example above, each of the 12 counties in Maryland missing the value of "Percent Physician Diversity," then, would be imputed at 10.22 percent.[23]

3.2 Risk Factors

Based on the literature review, Hilltop identified and operationalized approximately 190 risk factors to be included in the risk model. While some of these risk factors are eliminated in the variable selection step due to high inter-variable correlation, this process is data-driven, and all risk factors are included in the pool of *potential* risk factors to be used in the model. A high-level description of risk factors is provided in the Sections below. For a description of each risk factor, along with data source and sample statistics, see Appendix 1.

---

[22] For additional documentation, see https://usa.ipums.org/usa-action/variables/COUNTYFIP#description_section.
[23] Data are missing for counties in Puerto Rico, so these values are imputed at the national average.

3.2.1 Literature Review

As a first step in the development process for the Hilltop Pre-AH Model™, Hilltop conducted a comprehensive literature review. The goal of the review was to find peer-reviewed academic journal articles that identified risk factors for potentially avoidable hospital events, thus providing a basis for risk factor extraction and feature creation. Identified risk factors were coded using CCLF and other publicly available data sources and included in the final risk model as potential predictors of avoidable hospitalization or ED use. The literature review provided the foundation for the risk model and was a crucial step in the modeling process. Using inclusion and exclusion criteria designed to reflect the MDPCP patient population, the Hilltop team screened over 3,300 articles in both a primary and secondary literature search, ultimately selecting 211 articles for risk factor extraction. For additional detail, see Pelser et al. (2019).

3.2.2 Part A Risk Factors

Risk factors based on Part A claims cover information on admissions over the past 12 months; nursing home stays over the past 12 months; and certain procedures. Additionally, the Part A claims are used in order to construct the avoidable hospital event outcome, as well as the diagnostic condition flags. These condition flags rely on diagnostic information from Part A and Part B claims in conjunction with Chronic Conditions Data Warehouse (CCW) coding specifications in order to generate beneficiary-level risk factors that represent underlying disease states.[24]

3.2.3. Part B Risk Factors

Risk factors based on Part B claims cover utilization of certain services (such as vaccinations, lab tests, or J-code procedures), place of service (for example, urgent care or rural health clinic), and provider specialty (for example, endocrinology or oncology). Hilltop also created risk factors to capture a beneficiary's primary care utilization and continuity of care. Finally, as above, the Part B claims are used in order to construct the avoidable hospital event outcome, as well as the diagnostic condition flags.

3.2.4 Part D Risk Factors

Using Medicare Part D claims, Hilltop flags utilization of drugs identified in its literature review as potential risk factors for potentially avoidable hospital events. In order to capture compound drugs, which are drugs that contain multiple active ingredients, Hilltop relies largely on text-

---

[24] Additional detail on the CCW condition flag specifications can be found here: https://www.ccwdata.org/documents/10280/19139421/ccw-chronic-condition-algorithms.pdf, https://www.ccwdata.org/documents/10280/19139421/ccw-chronic-condition-algorithms-reference-list.pdf

The Hilltop Institute

based, "contains"-type searches of the FDA's "National Drug Code Directory."[25] See Table 2 below for Hilltop's primary search strategy, as well as for a list of the substances flagged.

Table 2. Primary Search Strategy for MDPCP Pharmacy Risk Factors

Risk Factor	Primary Search Method in NDC	Substances Flagged
Losartan use	Substance name contains "LOSARTAN"	LOSARTAN POTASSIUM; LOSARTAN POTASSIUM and HYDROCHLOROTHIAZIDE
Warfarin use	Substance name contains "WARFARIN"	WARFARIN SODIUM
Cilostazol use	Substance name contains "CILOSTAZOL"	CILOSTAZOL
Insulin use	Substance name or nonproprietary name contains "INSULIN" and marketing category name does not contain "UNAPPROVED"	INSULIN ASPART; INSULIN DEGLUDEC; INSULIN DEGLUDEC and LIRAGLUTIDE; INSULIN DETEMIR; INSULIN GLARGINE; INSULIN GLARGINE and LIXISENATIDE; INSULIN GLULISINE; INSULIN HUMAN; INSULIN LISPRO
Statin use	Drug Class contains "HMG-CoA Reductase Inhibitor" or substance name contains "ROSUVASTATIN CALCIUM"	SIMVASTATIN; LOVASTATIN; PITAVASTATIN; ROSUVASTATIN CALCIUM; PRAVASTATIN SODIUM; FLUVASTATIN SODIUM; PITAVASTATIN CALCIUM; ATORVASTATIN CALCIUM; ATORVASTATIN CALCIUM TRIHYDRATE; ATORVASTATIN CALCIUM PROPYLENE GLYCOL SOLVATE; EZETIMIBE and SIMVASTATIN; NIACIN and LOVASTATIN; SIMVASTATIN and NIACIN; AMLODIPINE BESYLATE and ATORVASTATIN CALCIUM; AMLODIPINE BESYLATE and ATORVASTATIN CALCIUM TRIHYDRATE
Leukotriene Receptor Modifier use	Drug class contains "Leukotriene Receptor Antagonist"	MONTELUKAST; MONTELUKAST SODIUM; ZAFIRLUKAST
Beta Blocker use	Substance Name contains "METOPROLOL" or "CARVEDILOL"[26]	CARVEDILOL; CARVEDILOL PHOSPHATE; METOPROLOL SUCCINATE; METOPROLOL TARTRATE; METOPROLOL TARTRATE and HYDROCHLOROTHIAZIDE; METOPROLOL SUCCINATE and HYDROCHLOROTHIAZIDE
Oral Corticosteroid use	Drug class contains "Corticosteroid" and route is "ORAL" and dosage form contains either "CAPSULE" or "TABLET" and marketing category does not contain "UNAPPROVED"	BUDESONIDE; CORTISONE ACETATE; DEFLAZACORT; DEXAMETHASONE; HYDROCORTISONE; METHYLPREDNISOLONE; PREDNISOLONE; PREDNISOLONE SODIUM PHOSPHATE
Antidiabetes Medication	Substance name contains "FLOZIN", "GLIPTIN", "THIAZOLIDINEDIONE", "ROSIGLITAZONE", or "PIOGLITAZONE"	ALOGLIPTIN BENZOATE; ALOGLIPTIN BENZOATE and METFORMIN HYDROCHLORIDE; ALOGLIPTIN BENZOATE and PIOGLITAZONE HYDROCHLORIDE; CANAGLIFLOZIN; CANAGLIFLOZIN and METFORMIN HYDROCHLORIDE; DAPAGLIFLOZIN; DAPAGLIFLOZIN PROPANEDIOL; DAPAGLIFLOZIN PROPANEDIOL and METFORMIN HYDROCHLORIDE; DAPAGLIFLOZIN and

---

[25] For example, "Simcor" contains two active substances: Simvastatin and Niacin. This is flagged as a statin because one of its active ingredients is a statin. Source for the FDA NDC directory: https://www.fda.gov/drugs/drug-approvals-and-databases/national-drug-code-directory

[26] Based on Table 1 from Brophy et al., 2001.

Risk Factor	Primary Search Method in NDC	Substances Flagged
		SAXAGLIPTIN HYDROCHLORIDE; EMPAGLIFLOZIN; EMPAGLIFLOZIN <u>and</u> LINAGLIPTIN; EMPAGLIFLOZIN <u>and</u> METFORMIN HYDROCHLORIDE; ERTUGLIFLOZIN PIDOLATE; ERTUGLIFLOZIN PIDOLATE <u>and</u> METFORMIN HYDROCHLORIDE; ERTUGLIFLOZIN PIDOLATE <u>and</u> SITAGLIPTIN PHOSPHATE; LINAGLIPTIN; LINAGLIPTIN <u>and</u> METFORMIN HYDROCHLORIDE; METFORMIN HYDROCHLORIDE <u>and</u> PIOGLITAZONE HYDROCHLORIDE; PIOGLITAZONE; PIOGLITAZONE HYDROCHLORIDE; PIOGLITAZONE HYDROCHLORIDE <u>and</u> GLIMEPIRIDE; PIOGLITAZONE HYDROCHLORIDE <u>and</u> METFORMIN HYDROCHLORIDE; ROSIGLITAZONE MALEATE; SAXAGLIPTIN HYDROCHLORIDE; SAXAGLIPTIN HYDROCHLORIDE <u>and</u> METFORMIN HYDROCHLORIDE; SITAGLIPTIN PHOSPHATE; SITAGLIPTIN PHOSPHATE <u>and</u> METFORMIN HYDROCHLORIDE
Oral Antibiotics	Substance name contains prescription names in NCQA "Antibiotics of Concern" and "All other Antibiotics" and route is "ORAL"– See Appendix 2.	See Appendix 2.

3.2.5 Environmental Risk Factors

Several of the risk factors Hilltop identified during the literature review were individual-level demographic and socioeconomic factors that are unavailable in the CCLF data (for example, marital status). Consequently, corresponding area-level risk factors (for example, the percentage of the population aged 15+ that is currently married) are included in the risk model in order to proxy for the unobserved individual-level variables. Other environmental risk factors (for example, the area poverty rate) are intended to capture the social determinants of health: the neighborhood conditions in which people live and age that may affect health outcomes. Hilltop plans to refine these measures to incorporate greater geographic granularity in future versions of the risk model and will update the documentation accordingly.

3.3 Modeling

Methodologically, Hilltop relies on a discrete time survival model that uses current values of procedural, diagnostic, utilization-based, pharmacy, demographic, and environmental risk factors to predict the likelihood that an individual incurs an avoidable hospitalization or ED visit in the *following* month. The parameter estimates generated in the model training are subsequently used to generate individual risk predictions in the scoring step. Each of these points are discussed below.

The Hilltop Institute

3.3.1 Avoidable Hospitalizations and Emergency Department Visits

The outcome measure in the Hilltop Pre-AH Model™ is a 0/1 indicator variable denoting whether an individual incurred an avoidable hospitalization or ED visit in a given month. In order to construct this measure, Hilltop relies on technical definitions provided by the Agency for Healthcare Research and Quality (AHRQ) as part of its prevention quality indicator (PQI) measures.[27] Diagnosis codes from Part A inpatient and ED claims are used to flag the following conditions, which are the basis for the composite PAH flag:[28]

- PQI #1: Diabetes Short-Term Complications
- PQI #3: Diabetes Long-Term Complications
- PQI #5: COPD or Asthma in Older Adults
- PQI #7: Hypertension
- PQI #8: Heart Failure
- PQI #10: Dehydration
- PQI #11: Bacterial Pneumonia
- PQI #12: Urinary Tract Infection
- PQI #14: Uncontrolled diabetes
- PQI #15: Asthma in Younger Adults
- PQI #16: Lower-Extremity Amputation among Patients with Diabetes

This is implemented in the model as an indicator variable at the person-month level. If an individual incurs at least one avoidable hospitalization or ED visit in a given month, then that person receives a value of 1 for this variable—and 0 otherwise.

3.3.2 Statistical Model

Avoidable hospitalization/ED visits are recurrent events with time-dependent covariates. Accordingly, the Hilltop Pre-AH Model™ is operationalized as a discrete-time survival model that uses the *current* month of risk factors in order to predict avoidable hospitalization/ED visits in the *following* month. The model uses month as a time unit—instead of, for example, weeks or years—in order to balance the increased model accuracy obtained by a more granular time unit with the increased prediction error due to rare events.

The raw CCLF data span three years, or 36 person-months for individuals with full coverage. Since the model estimates the risk of incurring an avoidable hospitalization in the *next* month,

---

[27] For more information, see https://www.qualityindicators.ahrq.gov/modules/pqi_resources.aspx.

[28] Specifically, Hilltop defines these claims as those with a claim type of either 60 or 61 (indicating an inpatient claim) or a claim type of 40 (indicating an outpatient claim) and revenue center codes of 0450-0459 and 0981. Source: https://www.resdac.org/articles/how-identify-hospital-claims-emergency-room-visits-medicare-claims-data.

however, it is not possible to use the most recently available month of risk data in the training model (since the next month's avoidable hospitalizations have not been realized at this point). Therefore, the training data is based on underlying data covering 35 person-months per attributed patient with full coverage. While, in general, a reduction in sample size can adversely impact the statistical precision of the risk factor estimates, lagged predictors are used for three reasons.

First, several of the risk factors—such as the count of hospitalizations in the previous 12 months, or the condition flag for diabetes—overlap with the definition of an avoidable hospital event. Consequently, including these risk factors as *contemporaneous* predictors could artificially increase the predictive power of the model. Second, Hilltop believes that using lagged predictors aids in the interpretability of the model. The goal of the Hilltop Pre-AH Model™ is to predict future avoidable hospitalizations, and using contemporaneous predictors to generate future risk scores requires the assumption that individuals' risk factors do not change in the future. Finally, the use of lagged predictors implies a natural "person-now" data set: the most recent month of risk factors, which is not included in the training data set.

The statistical model is trained on an 80 percent sample of our analytical person-month data set. The functional form of the statistical model is:

$$\log\left(\frac{p_i(t)}{1-p_i(t)}\right) = \varphi(t) + \beta X_i(t-1) + \Omega V_i$$

- $\varphi(t)$ is a cubic function of time, accounting for the time effect
- β and Ω are the vectors of model parameters to be determined by training data
- $X_i(t-1)$ is a vector of patient $i's$ time-dependent features in the previous month
- V_i is a vector of patient *i*'s time-independent features
- $p_i(t)$ is the probability of avoidable hospitalization or ED visit of patient i at time t (i.e., the month following the realization of the risk factors)
- t is duration of time in months
    - counting start from the first month of available data if the patient is in coverage longer than three years, or
    - counting from the coverage start month if the patient's coverage start is within three years

The statistical model uses six types of risk factors: diagnostic, pharmacy, procedural, utilization-based, geographic, and demographic. It is important to note that not all risk factors are available for every person-month. Hilltop uses a twelve-month lookback period for most of the time-varying risk factors, implying that an individual with, for example, five months of claims history will have incomplete information in her risk factors: if this individual truly has glaucoma, then it is possible that she will not amass the claims history by month five that meets the qualifications required for a glaucoma flag in our model. Furthermore, while most individuals in the data have

valid ZIP codes that link to the environmental risk factor data set, several hundred beneficiaries have ZIP codes for which there is no equivalent ZCTA, and therefore receive no environmental risk factors.[29] Table 3 presents the risk factor availability, depending on claims history and the availability of ZIP code data.

Table 3. Availability of Risk Factors for Scoring

		At Least 12 Months of Claims History	
		Yes	**No**
Availability of ZIP Code Risk Factors	Yes	Dx/Rx/Proc/Util/Geo/Demo	Geo/Demo
	No	Dx/Rx/Proc/Util/Demo	Demo

Risk factor availability is an issue for the "scoring" step, in which risk scores are assigned to every individual based on the parameter estimates derived in the training step. For example, suppose that the vector of estimated coefficients from the logistic regression is as follows in Table 4.

Table 4. Example Risk Model Coefficients

Risk Factor	Value for individual i
Asthma Flag	.1
…	
ZIP Code Income	-.00001
…	
Age	.02

These hypothetical risk factor coefficients indicate that, as expected, if an individual meets the clinical criteria for asthma, the risk of avoidable hospitalization is higher; if he or she lives in a ZIP code with higher income, the risk is lower; and if he or she is older, the risk is higher. The scoring step will apply this vector of coefficients to the "person-now;" that is, the current month for each individual. Individual *i*'s predicted probability of incurring an avoidable hospitalization in the next month, then, will be scored as follows:

$$Risk_i = \frac{e^{.1*Asthma_i+\cdots-.00001*ZIP\ Code\ Income_i+\cdots+.02*Age_i}}{1 + e^{.1*Asthma_i+\cdots-.00001*ZIP\ Code\ Income_i+\cdots+.02*Age_i}}$$

Suppose that these variables (Asthma Flag, ZIP Code Income, and Age) are the only three risk factors in the model. Furthermore, suppose that individual *i* has the following characteristics:

Table 5. Example Risk Factors

Risk Factor	Value for individual i
Asthma Flag	1
ZIP Code Income	$55,000
Age	66

---

[29] These individuals appear to use P.O. boxes as their mailing address, which, being point representations, do not have ZCTA areal equivalents.

Maryland
DEPARTMENT OF HEALTH

The Hilltop Institute

This hypothetical individual has asthma, lives in a ZIP Code in which the median income is $55,000, and is 72 years old. Then, that individual's risk of an avoidable hospital event in the following month is $\frac{e^{(.1*1 - .00001*55,000 + .02*66)}}{1 + e^{(.1*1 - .00001*55,000 + .02*66)}}$ = 70.47 percent.

Suppose, however, that this individual is newly eligible for Medicare and does not have sufficient claims history to meet the criteria for an asthma flag (anything under 12 months). In this instance, the individual might truly have asthma as an underlying disease state, but this is not observable. The individual's risk factors, then, are:

Table 6. Example Risk Factors with Missing Information

Risk Factor	Value for individual i
Asthma Flag	NOT OBSERVED
ZIP Code Income	$55,000
Age	66

If the model's coefficients are applied only to the risk factors that are *observed*, then this individual's estimated risk is 68.35 percent. By failing to account for the risk factors that are not present in the model, the risk of incurring an avoidable hospital event is underestimated for individual *i*.

Hilltop's solution to this issue is to estimate four different models based on the risk factors that are available for each group. This allows the risk factors that are present to "compensate," to a certain extent, for the risk factors that are missing due to data availability. For example, suppose that an individual lacks sufficient claims history to generate diagnostic risk factors but does have the following demographic risk factors: age, gender, and race. If gender is correlated with the unobserved diagnostic risk factors (if, for example, female beneficiaries are more likely to experience chronic conditions than male beneficiaries), then the coefficient for the "gender" risk factor will capture this correlation, and thus represent the marginal impact of being female *and* the portion of unobserved diagnostic risk factors that is correlated with gender. Consequently, if female beneficiaries are more likely to experience chronic conditions than male beneficiaries, then the risk factor coefficient for "gender" will be larger in the models without diagnostic risk factors than in the models with diagnostic risk factors.[30] By allowing observed risk factors to capture some of the predictive power of unobserved risk factors, the loss in predictive power due to missing data is minimized. Note that this method is analogous to that used in the CMS HCC Risk Adjustment Model (CMS, 2019, p. 80).

The four models are trained on the subset of person-months for which all risk factors are complete (that is, person-months with at least 12 months of claims history and a valid ZIP code), and include the following sets of risk factors (analogous to the four partitions of the person-month sample):

---

[30] This is, indeed, the case: in the full model, males are 7.4 percent less likely to experience an avoidable hospitalization than females. In the model with only demographic risk factors, however, males are 13.1 percent less likely to experience an avoidable hospitalization.

- ▪ Model 1: use Rx/Dx/Util/Proc/Geo/Dem risk factors

- ▪ Model 2: use Geo/Dem risk factors

- ▪ Model 3: use Rx/Dx/Util/Proc/Dem risk factors

- ▪ Model 4: use Dem risk factors

Variable selection can improve the performance of predictive models by reducing prediction variance and increasing generalizability (Bagherzadeh-Khiabani et al., 2016; Walter & Tiemeier, 2009). Hilltop performed this in three steps: first, the team selected initial risk factors based on an extensive literature review, which screened over 3,300 articles and ultimately selected 211 published, peer-reviewed papers from which to extract risk factors (see Section 3.2.1 for additional detail). This generated a pool of roughly 190 risk factors. Next, Hilltop implemented correlation analysis in order to eliminate redundant risk factors (implemented using PROC VARCLUS in SAS). Finally, Hilltop used stepwise selection in the multivariable logistic model in order to remove insignificant predictors from the model before adding significant predictors.

In the current version of the Hilltop Pre-AH Model™, the risk factors enter the model additively: that is, if an individual has both diabetes and heart failure diagnostic flags, then his or her risk score will reflect the risk coefficient for the diabetes flag, plus that of the heart failure flag. It is possible, however, that there is additional risk due to the fact of the beneficiary having *both* conditions, over and above the sum of the risks of having each condition. Future versions of the model will identify and include these interaction variables.

3.3.3 Scoring

The four risk models above are trained on the subset of data with at least twelve months of claims history and full ZIP code-level data in order to estimate the vectors of coefficients for the risk factors in each model. Section 4.1 below presents these coefficients for Model 1 as odds ratios. Then, using the most recently available month of risk factors (that is, the "person-now" data set), individuals are scored using the model coefficients that correspond to the risk factors available in the person-now data set.

The Hilltop Pre-AH Model™ will generate risk scores for the entire MDPCP cohort, but individual practices will receive only practice-specific risk scores. This has the consequence that, if a practice contains disproportionately high-risk patients, and another practice contains disproportionately low-risk patients, then the riskiest patients within each practice will differ in their *absolute* risk. This does not affect the interpretation of within-practice risk scores, but Hilltop will seek to incorporate additional information on absolute risk in future phases of the model. Figure 1 presents a diagram of this point.

Figure 1. Stratification of MDPCP Cohort by Practices/Physicians
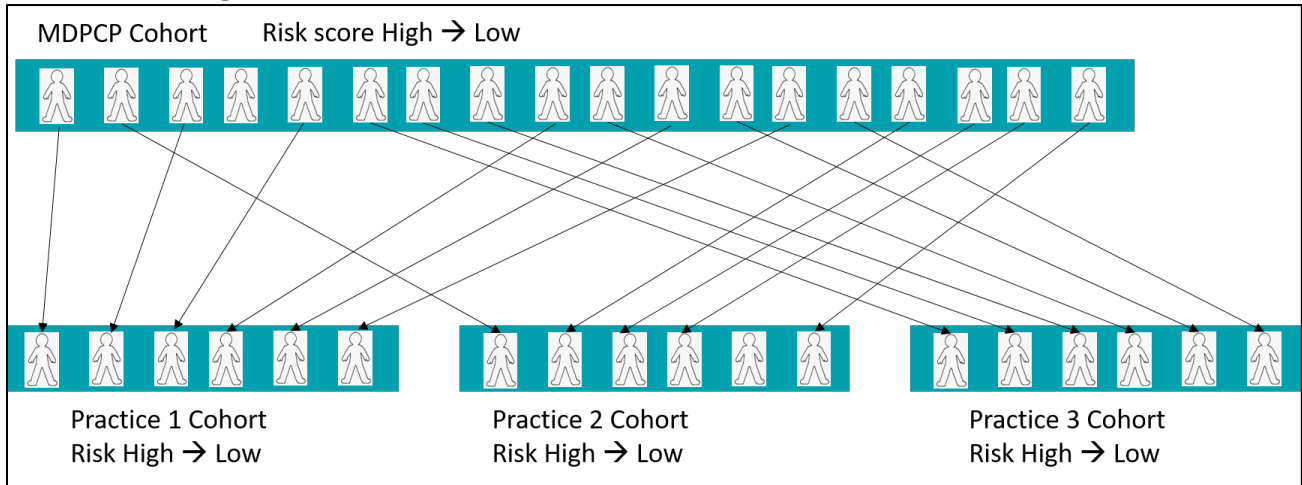
3.3.4 Model Training and Scoring Schedule

Hilltop plans to train the model once per quarter. This entails creating the risk factors from the raw claims data and estimating the four models discussed in Section 3.3.2, above. The resulting risk factor coefficients will then be used to score the MDPCP cohort once per *month*. This entails creating the risk factors from the raw claims data for the most recent month of claims history and then applying the most recent set of model coefficients.

For example, suppose that the four models are trained on January 1, 2019, April 1, 2019, July 1, 2019, and October 1, 2019, and that the previous month of claims data are available on the first day of the following month (so, in this example, claims data for June 2019 are available on July 1, 2019).[31] The model training generates risk factor coefficient estimates—one estimate for each risk factor—and these coefficient estimates are applied to the most recent set of risk factors in order to generate risk scores. For example, the coefficients estimated in the July 1, 2019 training will be used with the June 2019 risk factors in order to predict risk of avoidable hospitalization in July 2019 (which, remember, has not yet been observed as of the July 1, 2019 training date). These same risk factor coefficient estimates will be used with the July 2019 risk factors to predict August 2019 avoidable hospitalization and with August 2019 risk factors to predict September 2019 avoidable hospitalizations. Then, since the model is re-trained on October 1, 2019, September 2019 risk factors will be used with the new training model coefficients to predict avoidable hospitalizations in October 2019, and so on.

It is possible that the risk predictions may fall in accuracy as the training data model coefficients "fall behind" the person-now scoring data in time. For example, it is possible that, for the July 1, 2019 model training, the predictions are most accurate for July 2019 avoidable hospital events and then become less accurate for August 2019, and even less accurate for September 2019, since the training model coefficients become more removed from the current underlying data

---

[31] This example is only for the purposes of exposition; in reality, there is a lag of approximately 1 month in the CCLF claims.

generating process. Hilltop does not believe that this represents a significant threat to the model's predictive accuracy. Any systematic bias would have to be the result of underlying structural changes in the relationship of certain risk factors to the risk of incurring an avoidable hospital event, which seems unlikely to occur in a three-month window. However, Hilltop will monitor this and increase the training schedule as needed.

It is also possible that the latest month of claims for a given training data set will not be complete: for example, suppose that CCLF data received by Hilltop in July 2019 contain claims only through mid-June 2019. In this instance, Hilltop will still use these June 2019 risk factors in order to score the MDPCP cohort in July 2019, for two reasons. First, since all of the time-variant risk factors include at least 12 months of lookback, there is relatively low month-to-month variation; consequently, there is a relatively low chance of failing to include salient risk factors as a result of the missing data from the second half of June 2019. Second, it is imperative for the utility of the risk scores that the Hilltop Pre-AH Model™ uses the most recently available risk factors.

However, incomplete data for the final month of claims may introduce bias into the training model risk coefficients. To continue the example above, the training data will include beneficiaries' risk factors from May 2019, which are used to predict avoidable hospital events in June 2019. Since, in this example, claims information from late June 2019 will be missing, it will *appear* in the analytic data set as if most individuals did not incur an avoidable hospitalization in this month. Thus, in the extreme, May 2019 risk factors will be used to model all zeros for June 2019 outcomes, which would bias the risk factor coefficients toward zero. Given that 24 months of data are used in order to train the model, this is unlikely to **mean**ingfully affect the training coefficients. However, in order to prevent the possibility of bias, Hilltop will train the model only on the months of data for which we have full information for the outcome variable. To finish this example, May 2019 risk factors (and June 2019 outcomes) would not be included in the training model.

4. Model Performance

This section presents the details of the Hilltop Pre-AH Model™ performance using data from the period of February 2016 to January 2019.

4.1 Coefficients

Table 7 presents risk factor coefficient estimates for Model 1. This model includes all six types of risk factors: diagnostic, pharmacy, procedural, utilization-based, geographic, and demographic. The risk factors in Table 7 are those that were included in the final model. All other risk factors were eliminated in the variable selection steps due to either redundancy or insufficient predictive power. Note that the risk factor coefficients are presented as odds ratios, which can be interpreted in terms of a multiplicative impact: for example, an odds ratio of 1.05 indicates that if that risk factor were to increase by one unit, then the risk of incurring an avoidable hospitalization would increase by 5 percent.

Table 7. Risk Model Odds Ratios for Model 1

Risk Factor	Odds Ratio
Male (relative to female)	0.936
Race Unknown (relative to North American Native)	0.572
Race White (relative to North American Native)	0.772
Race Black (relative to North American Native)	1.063
Race Other (relative to North American Native)	0.686
Race Asian (relative to North American Native)	0.622
Race Hispanic (relative to North American Native)	0.714
Discharged to home (relative to transferred to post-acute care)	1.259
No prior hospitalizations (relative to transferred to post-acute care)	0.835
Other discharge status (relative to transferred to post-acute care)	2.367
Transferred to Inpatient (relative to transferred to post-acute care)	1.019
Number of avoidable hospitalizations	1.503
CCW indicator for chronic kidney disease	1.408
CCW indicator for Alzheimer's disease and related disorders or senile dementia	1.207
Indicator for urinary tract infection	1.349
CCW indicator for hypertension	1.278
CCW indicator for cataracts	0.858
CCW indicator for ischemic heart disease	1.271
CCW indicator for glaucoma	0.921
Indicator for peripheral and visceral atherosclerosis	1.125
CCW indicator for anemia	1.077
CCW indicator for atrial fibrillation	1.291
CCW indicator for osteoporosis	0.919
Indicator for problems with care provider dependency	1.335
Indicator for respiratory infection	1.125
CCW indicator for chronic obstructive pulmonary disease (COPD) and bronchiectasis	1.798
Indicator for pulmonary circulatory disorder	1.152
CCW indicator for asthma	1.243
CCW indicator for anxiety disorders	1.188
Indicator for neuropathy	1.215
CCW indicator for lung cancer	1.165
CCW indicator for mobility impairments	1.113
Indicator for pneumonia	1.131
CCW indicator for bipolar disorder	1.118
Indicator for lifestyle problems	1.221
CCW indicator for intellectual disabilities and related conditions	1.332
Indicator for albuminuria	1.113
CCW indicator for multiple sclerosis and transverse myelitis	1.334
Indicator for retinopathy	1.554
Number of prior admissions	0.922
Indicator for vaccination	0.872

Risk Factor	Odds Ratio
Indicator for previous conservative diabetic wound procedure	1.153
Number of primary care visits	1.004
Indicator for prior surgery	0.894
Number of urgent care visits	1.046
Indicator for provider administered drug	1.084
Continuity of primary care - Proportion	0.796
Indicator for oral antibiotic use	1.166
Indicator for statin use	0.854
Indicator for beta blocker use	1.122
Indicator for insulin use	1.262
National ranking of deprivation	1.005
Indicator for whole county mental health care shortage area	1.120
Indicator for dual Medicaid-Medicare enrollee	1.443
Indicator for hospice enrollment	1.407
Age in years	1.015

It is important to note that risk factor coefficient estimates will change as the model is re-trained. Risk factor coefficients for other models are available upon request.

4.2 Concentration Curves

Traditionally, the discriminatory power of predictive models has been summarized using the c-statistic, which is a measure of the area under the Receiver Operating Characteristic (ROC) curve (Steyerberg et al., 2016). The ROC curve plots the true positive rate against the false positive rate for binary classifiers using successive cutoff thresholds and "measures the probability that a randomly selected diseased subject has a higher predicted risk than a randomly selected non-diseased subject" (Maugen & Begg, 2016). However, this measure is uninformative regarding model calibration, which is the degree to which estimated risk scores match underlying true risk: it is possible to have a model with good discrimination and poor calibration (Alba et al., 2017). Moreover, the objective of the Hilltop Pre-AH Model™ is not binary classification, but instead the estimation of individual-level risks of incurring an avoidable hospital event so that care managers can, by focusing on the riskiest individuals, potentially intervene to prevent the most likely avoidable hospitalizations. To that end, the performance of the Hilltop Pre-AH Model™ is assessed using the *concentration curve*.[32]

This measure of model accuracy estimates the cumulative share of all avoidable hospital events incurred by the riskiest patients, where the reader can determine the share of all avoidable hospital events occurring for individuals above different risk thresholds. Below, find the concentration curve for patients in the hold-out sample in December 2018. In order to generate

---

[32] This is very similar to the Lorenz curve, which "is especially useful in the context of disease prevention because it maps out what public health policy investigators need to know. That is, it tells us how much disease burden will occur in any given proportion of the population with risks above a chosen threshold" (Maugen & Begg, 2016).

this curve, Hilltop trains the model using data from February 2016 to January 2019 on 80 percent of the MDPCP cohort and obtains the training model coefficients—like those listed in Table 7—using the method outlined above. Hilltop then applies these coefficients (as well as those from the other three groups as needed, depending on data availability) to the risk factors for the 20 percent of individuals *that have not been used to train the model* in the December 2018 MDPCP cohort. This step ensures that the model performance is not artificially inflated due to over-fitting of the estimation sample.

In order to estimate the concentration curve, the patient cohort is ordered from most to least risky (in terms of predicted risk) on the X axis, and the fraction of total avoidable hospital events captured by the riskiest patients on the Y axis. Note, for example, in Figure 2 that the top 10 percent riskiest patients account for approximately 50 percent of all avoidable hospitalizations in December 2018, and the top 20 percent riskiest patients account for approximately two-thirds of all avoidable hospitalizations. Figure 2 shows the curve for December 2018; concentration curves from other months are available upon request.

Figure 2. Concentration Curve as of December 2018
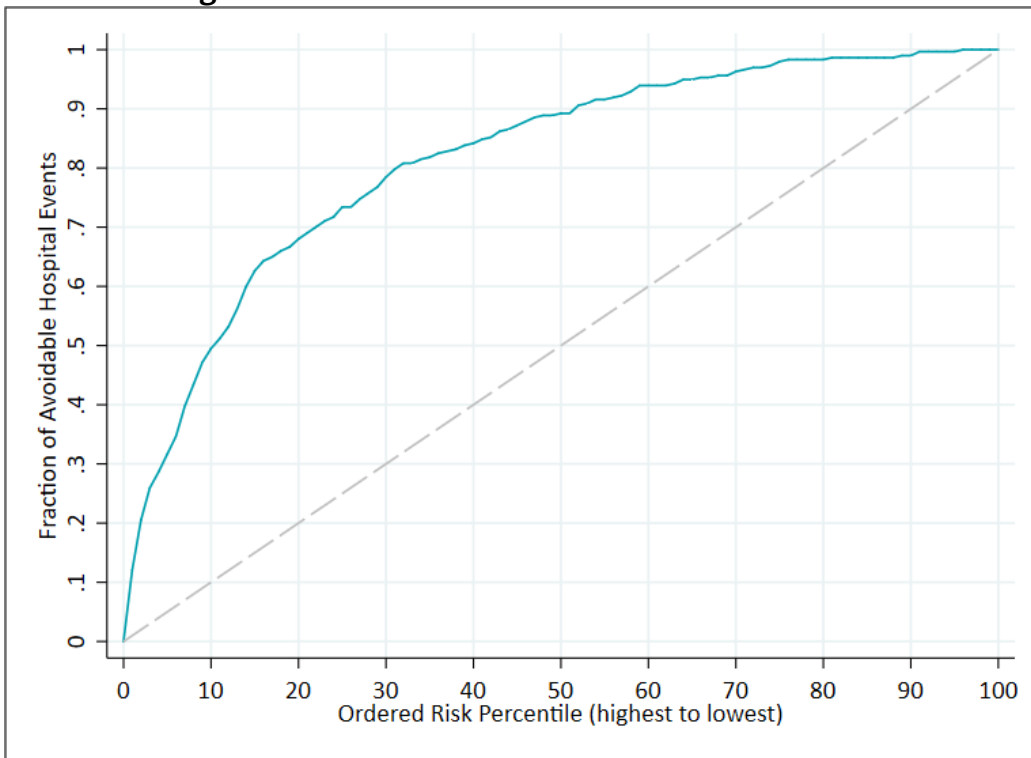
In order to assess the predictive power of the CMS HCC risk score, the same exercise is performed using that score instead of the Hilltop Pre-AH Model™ risk score. Figure 3 shows the resulting concentration curve. Note that the top 10 percent riskiest patients account for approximately 33 percent of all avoidable hospitalizations in December 2018, and the top 20 percent riskiest patients account for approximately 50 percent of all avoidable hospitalizations. Given a baseline of approximately 1,500 avoidable hospital events per month, this implies that if care managers were to rely solely on the CMS HCC risk score and focus on the riskiest 10 percent

of the cohort, then they would fail to identify 255 avoidable hospital events (relative to the number that would be identified using the Hilltop Pre-AH Model™ risk scores).

Figure 3. Concentration Curve for CMS HCC Risk Score as of December 2018
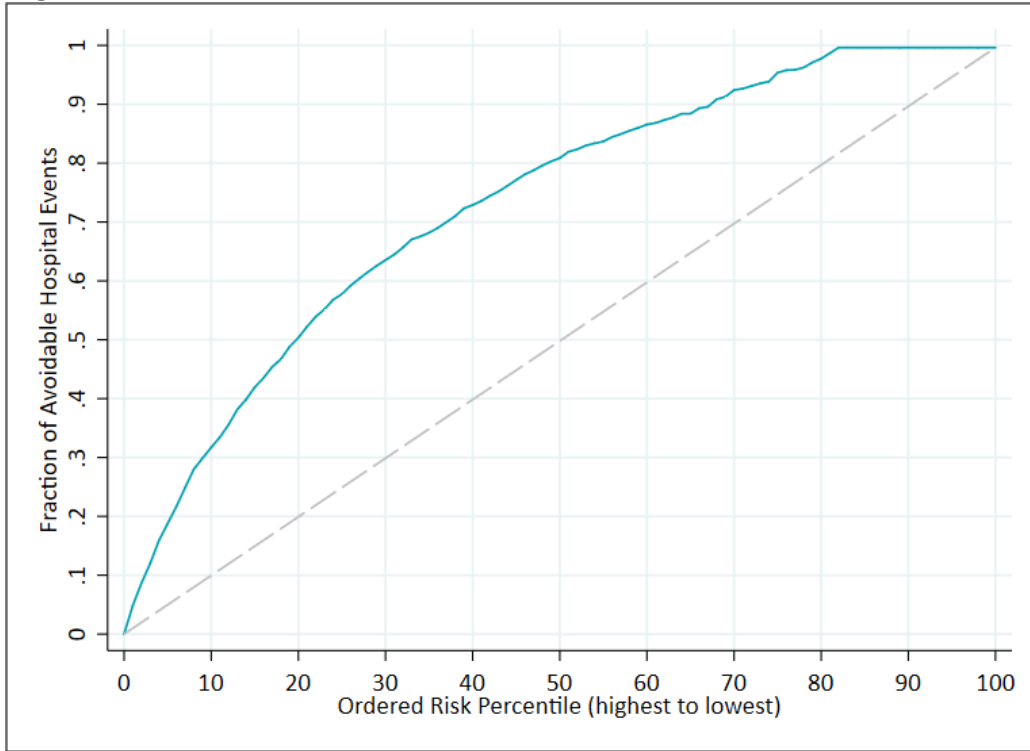
The Hilltop Pre-AH Model™ fit can be summarized using the Gini coefficient index, which is a measure (from 0 to 1) of the area between the concentration curve and the dotted 45-degree line. A higher Gini coefficient indicates greater model fit. Figure 4 plots the Gini coefficients from 23 months of data based on the concentration curves estimated on the 20 percent holdout sample from each month. The final month of holdout data – January 2019 – was not used due to incomplete information on avoidable hospital events in that month.

Figure 4. Gini Scores by Month
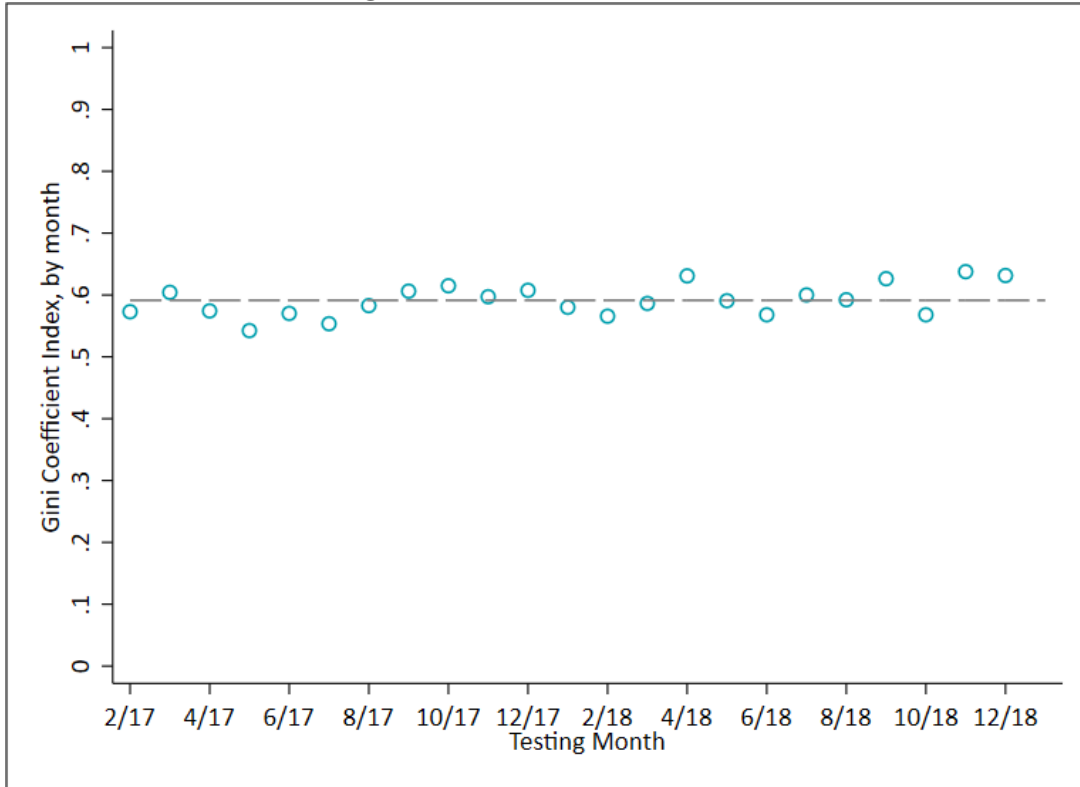
These scores indicate that model performance is steady over time, with a slight upward trend. This is consistent with the risk factors becoming more predictive as additional claims history becomes available.

4.3 Limitations

There are three main limitations of the Hilltop Pre-AH Model™ that are important to consider when implementing the model for guiding care coordination services. First, while the Medicare CCLF data on which many of the risk factors are based is the most timely data source available for this information, it does contain some temporal lag. With the expanded implementation of advanced revenue management systems, however, the difference in time between the date of service and the date of claim submission should continue to decrease. Second, it is likely that clinical information—such as patient vital signs and the results of laboratory tests—would increase the predictive power of the model but is not found in the claims data. Hilltop hopes to include this information in future versions of the model once the level of information exchange between electronic health records allows. Finally, even though ZIP code-level environmental risk factors are included, the variation inherent in each ZIP code introduces error. If available in the future, individual-level social welfare screening data will be added to provide a more robust individual-level risk prediction.

References

Alba, A. C., Agoritsas, T., Walsh, M., Hanna, S., Iorio, A., Devereaux, P. J., McGinn, T., & Guyatt, G. (2017.) Discrimination and calibration of clinical prediction models: Users' guides to the medical literature. *JAMA 318*(14)*,* 1377-1384.

Anderson, G. F., Ballreich, J., Bleich, S., Boyd, C., DuGoff, E., Leff, B., Salzburg, C., & Wolff, J. (2015). Attributes common to programs that successfully treat high-need, high-cost individuals. *The American Journal of Managed Care, 21*(11), e597-600.

Bagherzadeh-Khiabani, F., Ramezankhani, A., Azizi, F., Hadaegh, F., Steyerberg, E. W., & Khalili, D. (2016). A tutorial on variable selection for clinical prediction models: Feature selection methods in data mining could improve the results. *Journal of Clinical Epidemiology, 71*, 76-85.

Baker, J. M., Grant, R. W., & Gopalan, A. (2018). A systematic review of care management interventions targeting multimorbidity and high care utilization. *BMC Health Services Research, 18*(1), 65.

Berkowitz, S. A., Parashuram, S., Rowan, K., Andon, L., Bass, E. B., Bellantoni, M., Brotman, D. J., et al. (2018). Association of a care coordination model with health care costs and utilization: The Johns Hopkins Community Health Partnership (J-CHiP). *JAMA Network Open, 1*(7), e184273-e184273.

Billings, J., Zeitel, L., Lukomnik, J., Carey, T. S., Blank, A. E., & Newman, L. (1993). Impact of socioeconomic status on hospital use in New York City. *Health Affairs, 12*(1), 162-173.

Brophy, J. M., Joseph, L., & Rouleau, J. L. (2001). β-Blockers in congestive heart failure: A Bayesian meta-analysis. *Annals of Internal Medicine, 134*(7), 550-560.

Center for Medicare & Medicaid Innovation. (2017, December). *CPC+ payment methodologies: Beneficiary attribution, care management fee, performance-based incentive payment, and payment under the Medicare physician fee schedule. Version 1*. Retrieved from https://innovation.cms.gov/Files/x/cpcplus-methodology.pdf

Center for Medicare & Medicaid Innovation. (2019, April). *MDPCP payment methodologies: Beneficiary attribution, care management fee, performance-based incentive payment, and comprehensive care payment.* Version 1.0p.

Centers for Medicare & Medicaid Services. (2018, December). *Report to Congress: Risk adjustment in Medicare Advantage*. Retrieved from https://www.cms.gov/Medicare/Health-Plans/MedicareAdvtgSpecRateStats/Downloads/RTC-Dec2018.pdf

Congressional Research Service. (2006). *Changing postal ZIP code boundaries*. Retrieved from https://apps.dtic.mil/dtic/tr/fulltext/u2/a462043.pdf

Edwards, S. T., Peterson, K., Chan, B., Anderson, J., & Helfand, M. (2017). Effectiveness of intensive primary care interventions: A systematic review. *Journal of General Internal Medicine, 32*(12), 1377-1386.

Hong, C. S., Siegel, A. L., & Ferris, T. G. (2014). Caring for high-need, high-cost patients: What makes for a successful care management program. *Commonwealth Fund Issue Brief, 19*(9), 1-19.

Liaw, W., Moore, M., Iko, C., & Bazemore, A. (2015). Lessons for primary care from the first ten years of Medicare coordinated care demonstration projects. *J Am Board Fam Med, 28*(5), 556-564.

McCarthy, D., Ryan, J., & Klein, S. (2015). Models of care for high-need, high-cost patients: An evidence synthesis. *Commonwealth Fund Issue Brief, 31*, 1-19.

Mauguen, A., & Begg, C. B. (2016). Using the Lorenz curve to characterize risk predictiveness and etiologic heterogeneity. *Epidemiology, 27*(4), 531.

Peikes, D., Chen, A., Schore, J., & Brown, R. (2009). Effects of care coordination on hospitalization, quality of care, and health care expenditures among Medicare beneficiaries: 15 randomized trials. *JAMA, 301*(6), 603-618.

Pelser, C., Henderson, M., & Stockwell, I. (2019, April 23). *Risk factors for potentially avoidable hospital admissions and emergency department visits: A literature review.* Baltimore, MD: The Hilltop Institute, UMBC.

Ruggles, S., Flood, S., Goeken, R., Grover, J., Meyer, E., Pacas, J., & Sobek, M. (2019). *IPUMS USA: Version 9.0 [data set]*. Minneapolis, MN: IPUMS. https://doi.org/10.18128/D010.V9.0

Steyerberg, E. W., Vickers, A. J., Cook, N. R., Gerds, T., Gonen, M., Obuchowski, N., Pencina,M. J., & Kattan, M. W. (2010). Assessing the performance of prediction models: A framework for some traditional and novel measures. *Epidemiology, 21*(1), 128.

U.S. Census Bureau. (2018a). *TIGER/Line® Shapefiles. Technical documentation*. Retrieved from https://www2.census.gov/geo/pdfs/maps-data/data/tiger/tgrshp2018/TGRSHP2018_TechDoc.pdf

U.S. Census Bureau. (2018b). *ZIP Code Tabulation Areas (ZCTAs)*. Retrieved from https://www.census.gov/programs-surveys/geography/guidance/geo-areas/zctas.html

U.S. Postal Service Office of Inspector General. (2013). *The untold story of the ZIP code*. Report numbers RARC-WP-13-006. Retrieved from https://www.uspsoig.gov/sites/default/files/document-library-files/2015/rarc-wp-13-006_0.pdf

Walter, S., & Tiemeier, H. (2009). Variable selection: Current practice in epidemiological studies. *European Journal of Epidemiology, 24*(12), 733.

Appendix 1. Risk Factor Codebook

Beneficiary sex: Beneficiary's sex.

Source: Beneficiary Demographics

Risk Factor	Percentage
Male	.402
Female	.598

Beneficiary race: Beneficiary's Research Triangle Institute (RTI) race code.

Source: Beneficiary Demographics

Risk Factor	Percentage
White	.724
Black	.209
Hispanic	.009
Asian	.02
Native American	.001
Other	.017
Unknown	.02

Prior hospitalization discharge status: For each person-month, this variable indicates the discharge status of the person's most recently incurred inpatient hospital stay. If an individual has not incurred an inpatient hospital stay in the study period, he or she receives a value of "None."

Source: Part A claims

Risk Factor	Percentage
None	.89
Discharged to Home	.087
Other	.001
Transferred to Inpatient Care	.001
Transferred to Post-Acute Care	.022

Maryland
DEPARTMENT OF HEALTH

The Hilltop Institute

Prior hospitalization admission type: For each person-month, this variable indicates the admission type of the person's most recently incurred inpatient hospital stay. If an individual has not incurred an inpatient hospital stay in the study period, he or she receives a value of "None."

Source: Part A claims

Risk Factor	Percentage
Emergency	.07
Urgent	.005
Trauma Center	.001
Elective	.034
Newborn	0
Other	0
None	.89

Prior hospitalization admission source: For each person-month, this variable indicates the admission source of the person's most recently incurred inpatient hospital stay. If an individual has not incurred an inpatient hospital stay in the study period, he or she receives a value of "None."

Source: Part A claims

Risk Factor	Percentage
None	.89
Other	.002
Physician Referral	.098
Transfer from Facility	.01

Prior admission length of stay: For each person-month, this variable calculates the length of the most recently incurred hospital inpatient stay over the past 12 months. For individuals without a previous inpatient stay, this value is set to zero.

Source: Part A claims

Risk Factor	Mean	Minimum	Maximum
Prior admission length of stay	.484	0	181

Number of prior admissions: For each person-month, this variable counts the number of all inpatient hospital admissions incurred within the past twelve months.

Source: Part A claims

Risk Factor	Mean	Minimum	Maximum
Number of prior admissions	.162	0	22

Indicator for prior nursing home stay: For each person-month, this variable takes the value of 1 if a person incurred a nursing home stay within the last 12 months, and 0 otherwise.

Source: Part A claims

Risk Factor	Mean	Minimum	Maximum
Indicator for prior nursing home stay	.027	0	1

Number of heart-related procedures: For each person-month, this variable counts the number of heart-related procedures incurred over the past year.

Source: Part A claims

Risk Factor	Mean	Minimum	Maximum
Number of heart-related procedures	.019	0	14

Indicator for diabetic foot procedure: For each person-month, this variable takes the value of 1 if a person incurred an inpatient diabetic foot procedure over the last 12 months, and 0 otherwise.

Source: Part A claims

Risk Factor	Mean	Minimum	Maximum
Indicator for diabetic foot procedure	.001	0	1

Indicator for prior readmission: For each person-month, this variable takes the value of 1 if a person incurred an all-cause 30-day hospital readmission within the last 12 months, and 0 otherwise. Hilltop defines readmission as two inpatient stays occurring fewer than 30 days apart.

Source: Part A claims

Risk Factor	Mean	Minimum	Maximum
Indicator for prior readmission	.019	0	1

Maryland
DEPARTMENT OF HEALTH

The Hilltop Institute

Indicator for vaccination: For each person-month, this variable takes the value of 1 if a person received a vaccination within the past 12 months, and 0 otherwise.

Source: Part B claims

Risk Factor	Mean	Minimum	Maximum
Indicator for vaccination	.466	0	1

Number of HbA1c tests: For each person-month, this variable counts the number of visits within the past 12 months in which a person received a Hemoglobin A1C (HbA1c) test. Hilltop defines visits as unique combinations of person-provider-day.

Source: Part B claims

Risk Factor	Mean	Minimum	Maximum
Number of HbA1c tests	.551	0	16

Number of lab tests: For each person-month, this variable counts the number of visits within the past 12 months in which a person received any laboratory test. Hilltop defines visits as unique combinations of person-provider-day.

Source: Part B claims

Risk Factor	Mean	Minimum	Maximum
Number of lab tests	.147	0	33

Indicator for previous conservative diabetic wound procedure: For each person-month, this variable takes the value of 1 if a person underwent at least one conservative diabetic procedure within the past 12 months, and 0 otherwise.

Source: Part B claims

Risk Factor	Mean	Minimum	Maximum
Indicator for previous conservative diabetic wound procedure	.011	0	1

Number of primary care visits: For each person-month, this variable counts the number of primary care visits within the past 12 months. Hilltop defines visits as unique combinations of person-provider-day.

Source: Part B claims

Risk Factor	Mean	Minimum	Maximum
Number of primary care visits	8.869	0	213

Indicator for prior surgery: For each person-month, this variable takes the value of 1 if a person underwent a surgery within the past 12 months, and 0 otherwise.

Source: Part B claims

Risk Factor	Mean	Minimum	Maximum
Indicator for prior surgery	.567	0	1

Number of urgent care visits: For each person-month, this variable counts the number of urgent care visits incurred within the past 12 months. Hilltop defines visits as unique combinations of person-provider-day.

Source: Part B claims

Risk Factor	Mean	Minimum	Maximum
Number of urgent care visits	.138	0	42

Number of home health visits: For each person-month, this variable counts the number of home health visits incurred within the past 12 months. We apply a logarithmic transformation to non-zero values. Hilltop defines visits as unique combinations of person-provider-day.

Source: Part B claims

Risk Factor	Mean	Minimum	Maximum
Number of home health visits	.016	0	5.231106

Indicator for endocrinologist visit: For each person-month, this variable takes the value of 1 if a person visited an endocrinologist within the past 12 months, and 0 otherwise.

Source: Part B claims

Risk Factor	Mean	Minimum	Maximum
Indicator for endocrinologist visit	.067	0	1

Indicator for oncologist visit: For each person-month, this variable takes the value of 1 if a person visited an oncologist within the past 12 months, and 0 otherwise.

Source: Part B claims

Risk Factor	Mean	Minimum	Maximum
Indicator for oncologist visit	.092	0	1

Number of specialist visits: For each person-month, this variable counts the number of specialist visits incurred within the past 12 months. Hilltop defines visits as unique combinations of person-provider-day.

Source: Part B claims

Risk Factor	Mean	Minimum	Maximum
Number of specialist visits	4.986	0	365

Number of outpatient visits: For each person-month, this variable counts the number of visits in an outpatient setting incurred within the past 12 months. Hilltop defines visits as unique combinations of person-provider-day.

Source: Part B claims

Risk Factor	Mean	Minimum	Maximum
Number of outpatient visits	13.94	0	260

Number of rural clinic visits: For each person-month, this variable counts the number of rural clinic visits incurred within the past 12 months. Hilltop defines visits as unique combinations of person-provider-day.

Source: Part B claims

Risk Factor	Mean	Minimum	Maximum
Number of rural clinic visits	0	0	26

Continuity of primary care - Proportion: For each person-month, this variable estimates the fraction of primary care visits within the past 12 months provided by the same provider. For example, if a person had ten primary care visits over the past 12 months, and four visits were with the same provider, then this measure would take a value of .4. Hilltop defines visits as unique combinations of person-provider-day.

Source: Part B claims

Risk Factor	Mean	Minimum	Maximum
Continuity of primary care - Proportion	.466	0	1

Continuity of primary care - Index: For each person-month, this variable calculates the continuity of care index (Boxerman & Bice, 1977). This score ranges from 0 to 1 and is intended to measure dispersion in person-provider contact. If the person sees the same provider for all visits, the index score is 1; if the person sees a different physician for every visit, the index score

The Hilltop Institute

is 0. If a person has no primary care visits within the past year, they are assigned a value of 1. Hilltop defines visits as unique combinations of person-provider-day.

Source: Part B claims

Risk Factor	Mean	Minimum	Maximum
Continuity of primary care - Index	.386	0	1

Continuity of primary care - Duration: For each person-month, this variable calculates the average time interval between primary care visits over the past 12 months. Visits that occur within 14 days are aggregated. Individuals with no primary care visits over the past 12 months are assigned a value of 365. Hilltop defines visits as unique combinations of person-provider-day.

Source: Part B claims

Risk Factor	Mean	Minimum	Maximum
Continuity of primary care - Duration	129.513	15.20833	365

Indicator for mental health use: For each person-month, this variable takes the value of 1 if a person incurred a visit with a mental health professional over the past 12 months, and 0 otherwise.

Source: Part B claims

Risk Factor	Mean	Minimum	Maximum
Indicator for mental health use	.051	0	1

Indicator for provider administered drug: For each person-month, this variable takes the value of 1 if a person received a provider-administered drug as defined by a "J code" in the past 12 months, and 0 otherwise.

Source: Part B claims

Risk Factor	Mean	Minimum	Maximum
Indicator for provider administered drug	.21	0	1

Indicator for insulin use: For each person-month, this variable takes the value of 1 if a person incurred a claim for insulin within the past 12 months, and 0 otherwise.

Source: Part D claims

Risk Factor	Mean	Minimum	Maximum
Indicator for insulin use	.04	0	1

Indicator for anti-diabetes medication use: For each person-month, this variable takes the value of 1 if a person incurred a claim for anti-diabetes medication within the past 12 months, and 0 otherwise.

Source: Part D claims

Risk Factor	Mean	Minimum	Maximum
Indicator for anti-diabetes medication use	.033	0	1

Indicator for leukotriene receptor modifier use: For each person-month, this variable takes the value of 1 if a person incurred a claim for leukotriene receptor modifiers within the past 12 months, and 0 otherwise.

Source: Part D claims

Risk Factor	Mean	Minimum	Maximum
Indicator for leukotriene receptor modifier use	.029	0	1

Indicator for warfarin use: For each person-month, this variable takes the value of 1 if a person incurred a claim for warfarin within the past 12 months, and 0 otherwise.

Source: Part D claims

Risk Factor	Mean	Minimum	Maximum
Indicator for warfarin use	.028	0	1

Indicator for statin use: For each person-month, this variable takes the value of 1 if a person incurred a claim for statins within the past 12 months, and 0 otherwise.

Source: Part D claims

Risk Factor	Mean	Minimum	Maximum
Indicator for statin use	.345	0	1

Indicator for losartan use: For each person-month, this variable takes the value of 1 if a person incurred a claim for losartan within the past 12 months, and 0 otherwise.

Source: Part D claims

Risk Factor	Mean	Minimum	Maximum
Indicator for losartan use	.097	0	1

Maryland
DEPARTMENT OF HEALTH

The Hilltop Institute

Indicator for beta blocker use: For each person-month, this variable takes the value of 1 if a person incurred a claim for beta blockers within the past 12 months, and 0 otherwise.

Source: Part D claims

Risk Factor	Mean	Minimum	Maximum
Indicator for beta blocker use	.161	0	1

Indicator for cilostazol use: For each person-month, this variable takes the value of 1 if a person incurred a claim for cilostazol within the past 12 months, and 0 otherwise.

Source: Part D claims

Risk Factor	Mean	Minimum	Maximum
Indicator for cilostazol use	.002	0	1

Indicator for oral corticosteroid use: For each person-month, this variable takes the value of 1 if a person incurred a claim for oral corticosteroids within the past 12 months, and 0 otherwise.

Source: Part D claims

Risk Factor	Mean	Minimum	Maximum
Indicator for oral corticosteroid use	.051	0	1

Indicator for oral antibiotic use: For each person-month, this variable takes the value of 1 if a person incurred a claim for oral antibiotics within the past 12 months, and 0 otherwise.

Source: Part D claims

Risk Factor	Mean	Minimum	Maximum
Indicator for oral antibiotic use	.3	0	1

Number of medications: For each person-month, this variable counts the number of distinct medications (as measured by NDC codes) for which there are part D claims within the past 12 months.

Source: Part D claims

Risk Factor	Mean	Minimum	Maximum
Number of medications	6.734	0	146

Total health spending: For each person-month, this variable measures the total health spending incurred within the past 12 months. Hilltop defines this as the sum of claim total charge amount (Part A), claim payment amount (Part B claim lines, aggregated to the claim level), and claim line beneficiary payment amount (part D).

Source: Part A, B, and D claims

Risk Factor	Mean	Minimum	Maximum
Total health spending	10383.99	0	3372596

Number of avoidable hospitalizations: For each person-month, this variable counts the number of avoidable hospitalizations incurred within the prior 12 months (not including the month in which the avoidable hospitalization occurred).

Source: Part A claims

Risk Factor	Mean	Minimum	Maximum
Number of avoidable hospitalizations	.059	0	49

Avoidable hospitalization/ED visit: For each person-month, this variable records whether the individual incurred an avoidable hospitalization or ED visits in that month. We use the AHRQ's definition of avoidable hospitalization in defining this outcome. Please see Section 3.2.1 of this document for additional detail.

Source: Part A claims

Risk Factor	Mean	Minimum	Maximum
Avoidable hospitalization/ED visit	.006	0	1

Indicator for hospice enrollment: For each person-month, this variable takes the value of 1 if a person enrolled in hospice within the past 12 months, and 0 otherwise.

Source: Beneficiary Demographics

Risk Factor	Mean	Minimum	Maximum
Indicator for hospice enrollment	.001	0	1

Age: For each person-month, this variable records person age as of the end of the month.

Source: Beneficiary Demographics

Risk Factor	Mean	Minimum	Maximum
Age	72.58	20	110

Indicator for dual eligibility with Medicaid: For each person-month, this variable takes the value of 1 if a person was dually eligible for both Medicaid and Medicare within the past 12 months, and 0 otherwise.

Source: Beneficiary Demographics

Risk Factor	Mean	Minimum	Maximum
Indicator for dual eligibility with Medicaid	.132	0	1

CCW indicator for acquired hypothyroidism: For each person-month, this variable records whether the person meets the CCW clinical criteria for acquired hypothyroidism. If so, this variable takes the value of 1; if not, then 0.

Source: Part A and B claims

Risk Factor	Mean	Minimum	Maximum
CCW indicator for acquired hypothyroidism	.134	0	1

CCW indicator for acute myocardial infarction: For each person-month, this variable records whether the person meets the CCW clinical criteria for acute myocardial infarction. If so, this variable takes the value of 1; if not, then 0.

Source: Part A and B claims

Risk Factor	Mean	Minimum	Maximum
CCW indicator for acute myocardial infarction	.004	0	1

CCW indicator for anemia: For each person-month, this variable records whether the person meets the CCW clinical criteria for anemia. If so, this variable takes the value of 1; if not, then 0.

Source: Part A and B claims

Risk Factor	Mean	Minimum	Maximum
CCW indicator for anemia	.178	0	1

CCW indicator for asthma

CCW indicator for asthma: For each person-month, this variable records whether the person meets the CCW clinical criteria for asthma. If so, this variable takes the value of 1; if not, then 0.

Source: Part A and B claims

Risk Factor	Mean	Minimum	Maximum
CCW indicator for asthma	.048	0	1

CCW indicator for atrial fibrillation

CCW indicator for atrial fibrillation: For each person-month, this variable records whether the person meets the CCW clinical criteria for atrial fibrillation. If so, this variable takes the value of 1; if not, then 0.

Source: Part A and B claims

Risk Factor	Mean	Minimum	Maximum
CCW indicator for atrial fibrillation	.067	0	1

CCW indicator for benign prostatic hyperplasia

CCW indicator for benign prostatic hyperplasia: For each person-month, this variable records whether the person meets the CCW clinical criteria for benign prostatic hyperplasia. If so, this variable takes the value of 1; if not, then 0.

Source: Part A and B claims

Risk Factor	Mean	Minimum	Maximum
CCW indicator for benign prostatic hyperplasia	.063	0	1

CCW indicator for cataracts

CCW indicator for cataracts: For each person-month, this variable records whether the person meets the CCW clinical criteria for cataracts. If so, this variable takes the value of 1; if not, then 0.

Source: Part A and B claims

Risk Factor	Mean	Minimum	Maximum
CCW indicator for cataracts	.173	0	1

Maryland
DEPARTMENT OF HEALTH

The Hilltop Institute

CCW indicator for chronic kidney disease: For each person-month, this variable records whether the person meets the CCW clinical criteria for chronic kidney disease. If so, this variable takes the value of 1; if not, then 0.

Source: Part A and B claims

Risk Factor	Mean	Minimum	Maximum
CCW indicator for chronic kidney disease	.174	0	1

CCW indicator for chronic obstructive pulmonary disease (COPD) and bronchiectasis: For each person-month, this variable records whether the person meets the CCW clinical criteria for COPD and bronchiectasis. If so, this variable takes the value of 1; if not, then 0.

Source: Part A and B claims

Risk Factor	Mean	Minimum	Maximum
CCW indicator for chronic obstructive pulmonary disease (COPD) and bronchiectasis	.078	0	1

CCW indicator for diabetes: For each person-month, this variable records whether the person meets the CCW clinical criteria for diabetes. If so, this variable takes the value of 1; if not, then 0.

Source: Part A and B claims

Risk Factor	Mean	Minimum	Maximum
CCW indicator for diabetes	.254	0	1

CCW indicator for glaucoma: For each person-month, this variable records whether the person meets the CCW clinical criteria for glaucoma. If so, this variable takes the value of 1; if not, then 0.

Source: Part A and B claims

Risk Factor	Mean	Minimum	Maximum
CCW indicator for glaucoma	.125	0	1

The Hilltop Institute

CCW indicator for heart failure: For each person-month, this variable records whether the person meets the CCW clinical criteria for heart failure. If so, this variable takes the value of 1; if not, then 0.

Source: Part A and B claims

Risk Factor	Mean	Minimum	Maximum
CCW indicator for heart failure	.076	0	1

CCW indicator for hip/pelvic fracture: For each person-month, this variable records whether the person meets the CCW clinical criteria for hip/pelvic fracture. If so, this variable takes the value of 1; if not, then 0.

Source: Part A and B claims

Risk Factor	Mean	Minimum	Maximum
CCW indicator for hip/pelvic fracture	.004	0	1

CCW indicator for hyperlipidemia: For each person-month, this variable records whether the person meets the CCW clinical criteria for hyperlipidemia. If so, this variable takes the value of 1; if not, then 0.

Source: Part A and B claims

Risk Factor	Mean	Minimum	Maximum
CCW indicator for hyperlipidemia	.494	0	1

CCW indicator for hypertension: For each person-month, this variable records whether the person meets the CCW clinical criteria for hypertension. If so, this variable takes the value of 1; if not, then 0.

Source: Part A and B claims

Risk Factor	Mean	Minimum	Maximum
CCW indicator for hypertension	.567	0	1

The Hilltop Institute

CCW indicator for ischemic heart disease: For each person-month, this variable records whether the person meets the CCW clinical criteria for ischemic heart disease. If so, this variable takes the value of 1; if not, then 0.

Source: Part A and B claims

Risk Factor	Mean	Minimum	Maximum
CCW indicator for ischemic heart disease	.209	0	1

CCW indicator for osteoporosis: For each person-month, this variable records whether the person meets the CCW clinical criteria for osteoporosis. If so, this variable takes the value of 1; if not, then 0.

Source: Part A and B claims

Risk Factor	Mean	Minimum	Maximum
CCW indicator for osteoporosis	.056	0	1

CCW indicator for rheumatoid arthritis/osteoarthritis: For each person-month, this variable records whether the person meets the CCW clinical criteria for rheumatoid arthritis/osteoarthritis. If so, this variable takes the value of 1; if not, then 0.

Source: Part A and B claims

Risk Factor	Mean	Minimum	Maximum
CCW indicator for rheumatoid arthritis/osteoarthritis	.276	0	1

CCW indicator for female/male breast cancer: For each person-month, this variable records whether the person meets the CCW clinical criteria for female/male breast cancer. If so, this variable takes the value of 1; if not, then 0.

Source: Part A and B claims

Risk Factor	Mean	Minimum	Maximum
CCW indicator for female/male breast cancer	.034	0	1

CCW indicator for colorectal cancer

CCW indicator for colorectal cancer: For each person-month, this variable records whether the person meets the CCW clinical criteria for colorectal cancer. If so, this variable takes the value of 1; if not, then 0.

Source: Part A and B claims

Risk Factor	Mean	Minimum	Maximum
CCW indicator for colorectal cancer	.01	0	1

CCW indicator for prostate cancer: For each person-month, this variable records whether the person meets the CCW clinical criteria for prostate cancer. If so, this variable takes the value of 1; if not, then 0.

Source: Part A and B claims

Risk Factor	Mean	Minimum	Maximum
CCW indicator for prostate cancer	.032	0	1

CCW indicator for lung cancer: For each person-month, this variable records whether the person meets the CCW clinical criteria for lung cancer. If so, this variable takes the value of 1; if not, then 0.

Source: Part A and B claims

Risk Factor	Mean	Minimum	Maximum
CCW indicator for lung cancer	.007	0	1

CCW indicator for endometrial cancer: For each person-month, this variable records whether the person meets the CCW clinical criteria for endometrial cancer. If so, this variable takes the value of 1; if not, then 0.

Source: Part A and B claims

Risk Factor	Mean	Minimum	Maximum
CCW indicator for endometrial cancer	.003	0	1

CCW indicator for ADHD, conduct disorders, and hyperkinetic syndrome: For each person-month, this variable records whether the person meets the CCW clinical criteria for ADHD, conduct disorders, and hyperkinetic syndrome. If so, this variable takes the value of 1; if not, then 0.

Source: Part A and B claims

Risk Factor	Mean	Minimum	Maximum
CCW indicator for ADHD, conduct disorders, and hyperkinetic syndrome	.005	0	1

CCW indicator for alcohol use disorders: For each person-month, this variable records whether the person meets the CCW clinical criteria for alcohol use disorders. If so, this variable takes the value of 1; if not, then 0.

Source: Part A and B claims

Risk Factor	Mean	Minimum	Maximum
CCW indicator for alcohol use disorders	.002	0	1

CCW indicator for anxiety disorders: For each person-month, this variable records whether the person meets the CCW clinical criteria for anxiety disorders. If so, this variable takes the value of 1; if not, then 0.

Source: Part A and B claims

Risk Factor	Mean	Minimum	Maximum
CCW indicator for anxiety disorders	.103	0	1

CCW indicator for autism spectrum disorders: For each person-month, this variable records whether the person meets the CCW clinical criteria for autism spectrum disorders. If so, this variable takes the value of 1; if not, then 0.

Source: Part A and B claims

Risk Factor	Mean	Minimum	Maximum
CCW indicator for autism spectrum disorders	.001	0	1

CCW indicator for bipolar disorder: For each person-month, this variable records whether the person meets the CCW clinical criteria for bipolar disorder. If so, this variable takes the value of 1; if not, then 0.

Source: Part A and B claims

Risk Factor	Mean	Minimum	Maximum
CCW indicator for bipolar disorder	.019	0	1

CCW indicator for cerebral palsy: For each person-month, this variable records whether the person meets the CCW clinical criteria for cerebral palsy. If so, this variable takes the value of 1; if not, then 0.

Source: Part A and B claims

Risk Factor	Mean	Minimum	Maximum
CCW indicator for cerebral palsy	.002	0	1

CCW indicator for cystic fibrosis and other metabolic developmental disorders: For each person-month, this variable records whether the person meets the CCW clinical criteria for cystic fibrosis and other metabolic developmental disorders. If so, this variable takes the value of 1; if not, then 0.

Source: Part A and B claims

Risk Factor	Mean	Minimum	Maximum
CCW indicator for cystic fibrosis and other metabolic developmental disorders	.004	0	1

CCW indicator for drug use disorders: For each person-month, this variable records whether the person meets the CCW clinical criteria for drug use disorders. If so, this variable takes the value of 1; if not, then 0.

Source: Part A and B claims

Risk Factor	Mean	Minimum	Maximum
CCW indicator for drug use disorders	.002	0	1

The Hilltop Institute

CCW indicator for epilepsy: For each person-month, this variable records whether the person meets the CCW clinical criteria for epilepsy. If so, this variable takes the value of 1; if not, then 0.

Source: Part A and B claims

Risk Factor	Mean	Minimum	Maximum
CCW indicator for epilepsy	.016	0	1

CCW indicator for fibromyalgia, chronic pain and fatigue: For each person-month, this variable records whether the person meets the CCW clinical criteria for fibromyalgia, chronic pain and fatigue. If so, this variable takes the value of 1; if not, then 0.

Source: Part A and B claims

Risk Factor	Mean	Minimum	Maximum
CCW indicator for fibromyalgia, chronic pain and fatigue	.145	0	1

CCW indicator for intellectual disabilities and related conditions: For each person-month, this variable records whether the person meets the CCW clinical criteria for intellectual disabilities and related conditions. If so, this variable takes the value of 1; if not, then 0.

Source: Part A and B claims

Risk Factor	Mean	Minimum	Maximum
CCW indicator for intellectual disabilities and related conditions	.006	0	1

CCW indicator for learning disabilities: For each person-month, this variable records whether the person meets the CCW clinical criteria for learning disabilities. If so, this variable takes the value of 1; if not, then 0.

Source: Part A and B claims

Risk Factor	Mean	Minimum	Maximum
CCW indicator for learning disabilities	.001	0	1

The Hilltop Institute

CCW indicator for leukemias and lymphomas: For each person-month, this variable records whether the person meets the CCW clinical criteria for leukemias and lymphomas. If so, this variable takes the value of 1; if not, then 0.

Source: Part A and B claims

Risk Factor	Mean	Minimum	Maximum
CCW indicator for leukemias and lymphomas	.012	0	1

CCW indicator for liver disease, cirrhosis and other liver conditions (except viral hepatitis): For each person-month, this variable records whether the person meets the CCW clinical criteria for liver disease, cirrhosis and other liver conditions (except viral hepatitis). If so, this variable takes the value of 1; if not, then 0.

Source: Part A and B claims

Risk Factor	Mean	Minimum	Maximum
CCW indicator for liver disease, cirrhosis and other liver conditions (except viral hepatitis)	.026	0	1

CCW indicator for migraine and chronic headache: For each person-month, this variable records whether the person meets the CCW clinical criteria for migraine and chronic headache. If so, this variable takes the value of 1; if not, then 0.

Source: Part A and B claims

Risk Factor	Mean	Minimum	Maximum
CCW indicator for migraine and chronic headache	.021	0	1

CCW indicator for mobility impairments: For each person-month, this variable records whether the person meets the CCW clinical criteria for mobility impairments. If so, this variable takes the value of 1; if not, then 0.

Source: Part A and B claims

Risk Factor	Mean	Minimum	Maximum
CCW indicator for mobility impairments	.015	0	1

The Hilltop Institute

CCW indicator for multiple sclerosis and transverse myelitis: For each person-month, this variable records whether the person meets the CCW clinical criteria for multiple sclerosis and transverse myelitis. If so, this variable takes the value of 1; if not, then 0.

Source: Part A and B claims

Risk Factor	Mean	Minimum	Maximum
CCW indicator for multiple sclerosis and transverse myelitis	.005	0	1

CCW indicator for muscular dystrophy: For each person-month, this variable records whether the person meets the CCW clinical criteria for muscular dystrophy. If so, this variable takes the value of 1; if not, then 0.

Source: Part A and B claims

Risk Factor	Mean	Minimum	Maximum
CCW indicator for muscular dystrophy	0	0	1

CCW indicator for obesity: For each person-month, this variable records whether the person meets the CCW clinical criteria for obesity. If so, this variable takes the value of 1; if not, then 0.

Source: Part A and B claims

Risk Factor	Mean	Minimum	Maximum
CCW indicator for obesity	.124	0	1

CCW indicator for other developmental delays: For each person-month, this variable records whether the person meets the CCW clinical criteria for other developmental delays. If so, this variable takes the value of 1; if not, then 0.

Source: Part A and B claims

Risk Factor	Mean	Minimum	Maximum
CCW indicator for other developmental delays	.001	0	1

CCW indicator for peripheral vascular disease: For each person-month, this variable records whether the person meets the CCW clinical criteria for peripheral vascular disease. If so, this variable takes the value of 1; if not, then 0.

Source: Part A and B claims

Risk Factor	Mean	Minimum	Maximum
CCW indicator for peripheral vascular disease	.098	0	1

CCW indicator for personality disorders: For each person-month, this variable records whether the person meets the CCW clinical criteria for personality disorders. If so, this variable takes the value of 1; if not, then 0.

Source: Part A and B claims

Risk Factor	Mean	Minimum	Maximum
CCW indicator for personality disorders	.009	0	1

CCW indicator for post-traumatic stress disorder: For each person-month, this variable records whether the person meets the CCW clinical criteria for post-traumatic stress disorder. If so, this variable takes the value of 1; if not, then 0.

Source: Part A and B claims

Risk Factor	Mean	Minimum	Maximum
CCW indicator for post-traumatic stress disorder	.005	0	1

CCW indicator for pressure and chronic ulcers: For each person-month, this variable records whether the person meets the CCW clinical criteria for pressure and chronic ulcers. If so, this variable takes the value of 1; if not, then 0.

Source: Part A and B claims

Risk Factor	Mean	Minimum	Maximum
CCW indicator for pressure and chronic ulcers	.024	0	1

The Hilltop Institute

CCW indicator for sensory (blindness and visual) impairment: For each person-month, this variable records whether the person meets the CCW clinical criteria for sensory (blindness and visual) impairment. If so, this variable takes the value of 1; if not, then 0.

Source: Part A and B claims

Risk Factor	Mean	Minimum	Maximum
CCW indicator for sensory (blindness and visual) impairment	.003	0	1

CCW indicator for sensory (deafness and hearing) impairment: For each person-month, this variable records whether the person meets the CCW clinical criteria for sensory (deafness and hearing) impairment. If so, this variable takes the value of 1; if not, then 0.

Source: Part A and B claims

Risk Factor	Mean	Minimum	Maximum
CCW indicator for sensory (deafness and hearing) impairment	.044	0	1

CCW indicator for spina bifida and other congenital anomalies of the nervous system: For each person-month, this variable records whether the person meets the CCW clinical criteria for spina bifida and other congenital anomalies of the nervous system. If so, this variable takes the value of 1; if not, then 0.

Source: Part A and B claims

Risk Factor	Mean	Minimum	Maximum
CCW indicator for spina bifida and other congenital anomalies of the nervous system	.001	0	1

CCW indicator for spinal cord injury: For each person-month, this variable records whether the person meets the CCW clinical criteria for spinal cord injury. If so, this variable takes the value of 1; if not, then 0.

Source: Part A and B claims

Risk Factor	Mean	Minimum	Maximum
CCW indicator for spinal cord injury	.003	0	1

The Hilltop Institute

CCW indicator for tobacco use: For each person-month, this variable records whether the person meets the CCW clinical criteria for tobacco use. If so, this variable takes the value of 1; if not, then 0.

Source: Part A and B claims

Risk Factor	Mean	Minimum	Maximum
CCW indicator for tobacco use	.051	0	1

CCW indicator for traumatic brain injury and nonpsychotic mental disorders due to brain damage: For each person-month, this variable records whether the person meets the CCW clinical criteria for traumatic brain injury and nonpsychotic mental disorders due to brain damage. If so, this variable takes the value of 1; if not, then 0.

Source: Part A and B claims

Risk Factor	Mean	Minimum	Maximum
CCW indicator for traumatic brain injury and nonpsychotic mental disorders due to brain damage	.002	0	1

CCW indicator for viral hepatitis: For each person-month, this variable records whether the person meets the CCW clinical criteria for viral hepatitis. If so, this variable takes the value of 1; if not, then 0.

Source: Part A and B claims

Risk Factor	Mean	Minimum	Maximum
CCW indicator for viral hepatitis	.009	0	1

Indicator for arrhythmia: For each person-month, this variable records whether the person has incurred at least one inpatient or two non-inpatient claims with any diagnosis for arrhythmia within the past two years. If so, this variable takes the value of 1; if not, then 0.

Source: Part A and B claims

Risk Factor	Mean	Minimum	Maximum
Indicator for arrhythmia	.157	0	1

Indicator for albuminuria: For each person-month, this variable records whether the person has incurred at least one inpatient or two non-inpatient claims with any diagnosis for albuminuria within the past two years. If so, this variable takes the value of 1; if not, then 0.

Source: Part A and B claims

Risk Factor	Mean	Minimum	Maximum
Indicator for albuminuria	.014	0	1

Indicator for peptic ulcer disease: For each person-month, this variable records whether the person has incurred at least one inpatient or two non-inpatient claims with any diagnosis for peptic ulcer disease within the past two years. If so, this variable takes the value of 1; if not, then 0.

Source: Part A and B claims

Risk Factor	Mean	Minimum	Maximum
Indicator for peptic ulcer disease	.008	0	1

Indicator for cerebrovascular disease: For each person-month, this variable records whether the person has incurred at least one inpatient or two non-inpatient claims with any diagnosis for cerebrovascular disease within the past two years. If so, this variable takes the value of 1; if not, then 0.

Source: Part A and B claims

Risk Factor	Mean	Minimum	Maximum
Indicator for cerebrovascular disease	.089	0	1

Indicator for diabetes with complications: For each person-month, this variable records whether the person has incurred at least one inpatient or two non-inpatient claims with any diagnosis for diabetes with complications within the past two years. If so, this variable takes the value of 1; if not, then 0.

Source: Part A and B claims

Risk Factor	Mean	Minimum	Maximum
Indicator for diabetes with complications	.142	0	1

Indicator for fluid and electrolyte imbalance: For each person-month, this variable records whether the person has incurred at least one inpatient or two non-inpatient claims with any diagnosis for fluid and electrolyte imbalance within the past two years. If so, this variable takes the value of 1; if not, then 0.

Source: Part A and B claims

Risk Factor	Mean	Minimum	Maximum
Indicator for fluid and electrolyte imbalance	.097	0	1

Indicator for rheumatoid arthritis/collagen vascular disease: For each person-month, this variable records whether the person has incurred at least one inpatient or two non-inpatient claims with any diagnosis for rheumatoid arthritis/collagen vascular disease within the past two years. If so, this variable takes the value of 1; if not, then 0.

Source: Part A and B claims

Risk Factor	Mean	Minimum	Maximum
Indicator for rheumatoid arthritis/collagen vascular disease	.046	0	1

Indicator for metastatic cancer: For each person-month, this variable records whether the person has incurred at least one inpatient or two non-inpatient claims with any diagnosis for metastatic cancer within the past two years. If so, this variable takes the value of 1; if not, then 0.

Source: Part A and B claims

Risk Factor	Mean	Minimum	Maximum
Indicator for metastatic cancer	.01	0	1

Indicator for solid tumor without metastasis: For each person-month, this variable records whether the person has incurred at least one inpatient or two non-inpatient claims with any diagnosis for solid tumor without metastasis within the past two years. If so, this variable takes the value of 1; if not, then 0.

Source: Part A and B claims

Risk Factor	Mean	Minimum	Maximum
Indicator for solid tumor without metastasis	.092	0	1

Indicator for pulmonary circulatory disorder: For each person-month, this variable records whether the person has incurred at least one inpatient or two non-inpatient claims with any diagnosis for pulmonary circulatory disorder within the past two years. If so, this variable takes the value of 1; if not, then 0.

Source: Part A and B claims

Risk Factor	Mean	Minimum	Maximum
Indicator for pulmonary circulatory disorder	.028	0	1

Indicator for gastroesophageal reflux disease: For each person-month, this variable records whether the person has incurred at least one inpatient or two non-inpatient claims with any diagnosis for gastroesophageal reflux disease within the past two years. If so, this variable takes the value of 1; if not, then 0.

Source: Part A and B claims

Risk Factor	Mean	Minimum	Maximum
Indicator for gastroesophageal reflux disease	.167	0	1

Indicator for gastroparesis: For each person-month, this variable records whether the person has incurred at least one inpatient or two non-inpatient claims with any diagnosis for gastroparesis within the past two years. If so, this variable takes the value of 1; if not, then 0.

Source: Part A and B claims

Risk Factor	Mean	Minimum	Maximum
Indicator for gastroparesis	.003	0	1

Indicator for protein-calorie malnutrition: For each person-month, this variable records whether the person has incurred at least one inpatient or two non-inpatient claims with any diagnosis for protein-calorie malnutrition within the past two years. If so, this variable takes the value of 1; if not, then 0.

Source: Part A and B claims

Risk Factor	Mean	Minimum	Maximum
Indicator for protein-calorie malnutrition	.005	0	1

Indicator for sleep apnea: For each person-month, this variable records whether the person has incurred at least one inpatient or two non-inpatient claims with any diagnosis for sleep apnea within the past two years. If so, this variable takes the value of 1; if not, then 0.

Source: Part A and B claims

Risk Factor	Mean	Minimum	Maximum
Indicator for sleep apnea	.099	0	1

Indicator for diabetic ulcer: For each person-month, this variable records whether the person has incurred at least one inpatient or two non-inpatient claims with any diagnosis for diabetic ulcer within the past two years. If so, this variable takes the value of 1; if not, then 0.

Source: Part A and B claims

Risk Factor	Mean	Minimum	Maximum
Indicator for diabetic ulcer	.033	0	1

Indicator for urinary tract infection: For each person-month, this variable records whether the person has incurred at least one inpatient or two non-inpatient claims with any diagnosis for urinary tract infection within the past two years. If so, this variable takes the value of 1; if not, then 0.

Source: Part A and B claims

Risk Factor	Mean	Minimum	Maximum
Indicator for urinary tract infection	.086	0	1

Indicator for sepsis: For each person-month, this variable records whether the person has incurred at least one inpatient or two non-inpatient claims with any diagnosis for sepsis within the past two years. If so, this variable takes the value of 1; if not, then 0.

Source: Part A and B claims

Risk Factor	Mean	Minimum	Maximum
Indicator for sepsis	.018	0	1

The Hilltop Institute

Indicator for neuropathy: For each person-month, this variable records whether the person has incurred at least one inpatient or two non-inpatient claims with any diagnosis for neuropathy within the past two years. If so, this variable takes the value of 1; if not, then 0.

Source: Part A and B claims

Risk Factor	Mean	Minimum	Maximum
Indicator for neuropathy	.048	0	1

Indicator for retinopathy: For each person-month, this variable records whether the person has incurred at least one inpatient or two non-inpatient claims with any diagnosis for retinopathy within the past two years. If so, this variable takes the value of 1; if not, then 0.

Source: Part A and B claims

Risk Factor	Mean	Minimum	Maximum
Indicator for retinopathy	.002	0	1

Indicator for problems with education and literacy: For each person-month, this variable records whether the person has incurred at least one inpatient or two non-inpatient claims with any diagnosis for problems with education and literacy within the past two years. If so, this variable takes the value of 1; if not, then 0.

Source: Part A and B claims

Risk Factor	Mean	Minimum	Maximum
Indicator for problems with education and literacy	0	0	1

Indicator for problems with employment and unemployment: For each person-month, this variable records whether the person has incurred at least one inpatient or two non-inpatient claims with any diagnosis for problems with employment and unemployment within the past two years. If so, this variable takes the value of 1; if not, then 0.

Source: Part A and B claims

Risk Factor	Mean	Minimum	Maximum
Indicator for problems with employment and unemployment	0	0	1

Indicator for occupational exposure to risk factors: For each person-month, this variable records whether the person has incurred at least one inpatient or two non-inpatient claims with any diagnosis for occupational exposure to risk factors within the past two years. If so, this variable takes the value of 1; if not, then 0.

Source: Part A and B claims

Risk Factor	Mean	Minimum	Maximum
Indicator for occupational exposure to risk factors	0	0	1

Indicator for problems with housing and economic conditions: For each person-month, this variable records whether the person has incurred at least one inpatient or two non-inpatient claims with any diagnosis for problems with housing and economic conditions within the past two years. If so, this variable takes the value of 1; if not, then 0.

Source: Part A and B claims

Risk Factor	Mean	Minimum	Maximum
Indicator for problems with housing and economic conditions	.001	0	1

Indicator for problems with social environment: For each person-month, this variable records whether the person has incurred at least one inpatient or two non-inpatient claims with any diagnosis for problems with social environment within the past two years. If so, this variable takes the value of 1; if not, then 0.

Source: Part A and B claims

Risk Factor	Mean	Minimum	Maximum
Indicator for problems with social environment	.002	0	1

Indicator for problems with upbringing: For each person-month, this variable records whether the person has incurred at least one inpatient or two non-inpatient claims with any diagnosis for problems with upbringing within the past two years. If so, this variable takes the value of 1; if not, then 0.

Source: Part A and B claims

Risk Factor	Mean	Minimum	Maximum
Indicator for problems with upbringing	0	0	1

Indicator for other problems with primary support group: For each person-month, this variable records whether the person has incurred at least one inpatient or two non-inpatient claims with any diagnosis for other problems with primary support group within the past two years. If so, this variable takes the value of 1; if not, then 0.

Source: Part A and B claims

Risk Factor	Mean	Minimum	Maximum
Indicator for other problems with primary support group	.001	0	1

Indicator for psychosocial problems: For each person-month, this variable records whether the person has incurred at least one inpatient or two non-inpatient claims with any diagnosis for psychosocial problems within the past two years. If so, this variable takes the value of 1; if not, then 0.

Source: Part A and B claims

Risk Factor	Mean	Minimum	Maximum
Indicator for psychosocial problems	0	0	1

Indicator for lifestyle problems: For each person-month, this variable records whether the person has incurred at least one inpatient or two non-inpatient claims with any diagnosis for lifestyle problems within the past two years. If so, this variable takes the value of 1; if not, then 0.

Source: Part A and B claims

Risk Factor	Mean	Minimum	Maximum
Indicator for lifestyle problems	.019	0	1

Indicator for difficulty with life management: For each person-month, this variable records whether the person has incurred at least one inpatient or two non-inpatient claims with any diagnosis for difficulty with life management within the past two years. If so, this variable takes the value of 1; if not, then 0.

Source: Part A and B claims

Risk Factor	Mean	Minimum	Maximum
Indicator for difficulty with life management	.001	0	1

Indicator for problems with care provider dependency: For each person-month, this variable records whether the person has incurred at least one inpatient or two non-inpatient claims with any diagnosis for problems with care provider dependency within the past two years. If so, this variable takes the value of 1; if not, then 0.

Source: Part A and B claims

Risk Factor	Mean	Minimum	Maximum
Indicator for problems with care provider dependency	.044	0	1

Indicator for pneumonia: For each person-month, this variable records whether the person has incurred at least one inpatient or two non-inpatient claims with any diagnosis for pneumonia within the past two years. If so, this variable takes the value of 1; if not, then 0.

Source: Part A and B claims

Risk Factor	Mean	Minimum	Maximum
Indicator for pneumonia	.034	0	1

Indicator for pancreatitis: For each person-month, this variable records whether the person has incurred at least one inpatient or two non-inpatient claims with any diagnosis for pancreatitis within the past two years. If so, this variable takes the value of 1; if not, then 0.

Source: Part A and B claims

Risk Factor	Mean	Minimum	Maximum
Indicator for pancreatitis	.009	0	1

Indicator for respiratory infection: For each person-month, this variable records whether the person has incurred at least one inpatient or two non-inpatient claims with any diagnosis for respiratory infection within the past two years. If so, this variable takes the value of 1; if not, then 0.

Source: Part A and B claims

Risk Factor	Mean	Minimum	Maximum
Indicator for respiratory infection	.123	0	1

Indicator for peripheral and visceral atherosclerosis: For each person-month, this variable records whether the person has incurred at least one inpatient or two non-inpatient claims with any diagnosis for peripheral and visceral atherosclerosis within the past two years. If so, this variable takes the value of 1; if not, then 0.

Source: Part A and B claims

Risk Factor	Mean	Minimum	Maximum
Indicator for peripheral and visceral atherosclerosis	.084	0	1

CCW indicator for depression and depressive disorders: For each person-month, this variable records whether the person meets the CCW clinical criteria for depression and depressive disorders. If so, this variable takes the value of 1; if not, then 0.

Source: Part A and B claims

Risk Factor	Mean	Minimum	Maximum
CCW indicator for depression and depressive disorders	.114	0	1

CCW indicator for Alzheimer's disease and related disorders or senile dementia: For each person-month, this variable records whether the person meets the CCW clinical criteria for Alzheimer's disease and related disorders or senile dementia. If so, this variable takes the value of 1; if not, then 0.

Source: Part A and B claims

Risk Factor	Mean	Minimum	Maximum
CCW indicator for Alzheimer's disease and related disorders or senile dementia	.054	0	1

CCW indicator for HIV/AIDS: For each person-month, this variable records whether the person meets the CCW clinical criteria for HIV/AIDS. If so, this variable takes the value of 1; if not, then 0.

Source: Part A and B claims

Risk Factor	Mean	Minimum	Maximum
CCW indicator for HIV/AIDS	.003	0	1

CCW indicator for stroke/ischemic transient attack: For each person-month, this variable records whether the person meets the CCW clinical criteria for stroke/ischemic transient attack. If so, this variable takes the value of 1; if not, then 0.

Source: Part A and B claims

Risk Factor	Mean	Minimum	Maximum
CCW indicator for stroke/ischemic transient attack	.032	0	1

CCW indicator for schizophrenia and other psychotic disorders: For each person-month, this variable records whether the person meets the CCW clinical criteria for schizophrenia and other psychotic disorders. If so, this variable takes the value of 1; if not, then 0.

Source: Part A and B claims

Risk Factor	Mean	Minimum	Maximum
CCW indicator for schizophrenia and other psychotic disorders	.012	0	1

Median household income: For each person, this variable records the median household income in the person's ZIP code tabulation area of residence (pooled from 2013-2017).

Source: American Community Survey (2017, 5-year estimates)

Risk Factor	Mean	Minimum	Maximum
Median household income	82047.05	14128	218638

Percentage in poverty: For each person, this variable records the percentage of families whose income in the past 12 months is below the poverty level in the person's ZIP code tabulation area of residence (pooled from 2013-2017) .

Source: American Community Survey (2017, 5-year estimates)

Risk Factor	Mean	Minimum	Maximum
Percentage in poverty	6.908	0	56.5

Percentage in poverty age 65+: For each person, this variable records the percentage of people age 65+ whose income in the past 12 months is below the poverty level in the person's ZIP code tabulation area of residence (pooled from 2013-2017).

Source: American Community Survey (2017, 5-year estimates)

Risk Factor	Mean	Minimum	Maximum
Percentage in poverty age 65+	7.873	0	88.9

Taxable interest per capita: For each person, this variable records taxable interest (tax year 2016) per person in the person's ZIP code tabulation area of residence.

Source: IRS Statistics of Income and American Community Survey (2017, 5-year estimates)

Risk Factor	Mean	Minimum	Maximum
Taxable interest per capita	223.889	0	8611.111

Percentage with less than high school education: For each person, this variable records the percent of individuals age 18 and older with less than a high school diploma in the person's ZIP code tabulation area of residence.

Source: American Community Survey (2017, 5-year estimates)

Risk Factor	Mean	Minimum	Maximum
Percentage with less than high school education	9.865	0	41.31013

Percentage foreign born: For each person, this variable records the percent of individuals who are foreign-born in the person's ZIP code tabulation area of residence.

Source: American Community Survey (2017, 5-year estimates)

Risk Factor	Mean	Minimum	Maximum
Percentage foreign born	13.2	0	74.63286

Percentage non-English speakers: For each person, this variable records the percent of individuals who speak Spanish or other languages and who speak English less than 'very well' in the person's ZIP code tabulation area of residence.

Source: American Community Survey (2017, 5-year estimates)

Risk Factor	Mean	Minimum	Maximum
Percentage non-English speakers	5.797	0	58.3195

Rurality index: For each person, this variable records the rural/urban index for the person's ZIP code tabulation area of residence. These data are composed of 10 codes that "delineate metropolitan, micropolitan, small town, and rural commuting areas based on the size and direction of the primary (largest) commuting flows." Higher values indicate a greater degree of rurality.

Source: USDA Rural-Urban Commuting Area Codes

Risk Factor	Mean	Minimum	Maximum
Rurality index	1.361	0	10

National ranking of deprivation: For each person, this variable records the national ranking of deprivation for the person's ZIP code tabulation area of residence. This index "includes factors for the theoretical domains of income, education, employment, and housing quality." See https://www.neighborhoodatlas.medicine.wisc.edu/ for additional detail. Higher values indicate a greater degree of deprivation.

Source: Neighborhood Atlas

Risk Factor	Mean	Minimum	Maximum
National ranking of deprivation	30.094	0	95

Indicator for presence of mental health center: For each person, this variable records whether the person's ZIP code tabulation area of residence contains at least one active community mental health center.

Source: CMS Provider of Service Files (December 2018)

Risk Factor	Mean	Minimum	Maximum
Indicator for presence of mental health center	.023	0	1

Indicator for presence of federally qualified health center: For each person, this variable records whether the person's ZIP code tabulation area of residence contains at least one active federally qualified health center.

Source: CMS Provider of Service Files (December 2018)

Risk Factor	Mean	Minimum	Maximum
Indicator for presence of federally qualified health center	.271	0	1

Indicator for presence of rural health clinic: For each person, this variable records whether the person's ZIP code tabulation area of residence contains at least one active rural health clinic.

Source: CMS Provider of Service Files (December 2018)

Risk Factor	Mean	Minimum	Maximum
Indicator for presence of rural health clinic	0	0	1

Indicator for presence of VA clinic or VA medical center: For each person, this variable records whether the person's ZIP code tabulation area of residence contains at least one VA clinic or medical center.

Source: Veterans Affairs Facility Listing

Risk Factor	Mean	Minimum	Maximum
Indicator for presence of VA clinic or VA medical center	.069	0	1

Indicator for presence of a for-profit hospital: For each person, this variable records whether the person's ZIP code tabulation area of residence contains at least one active (short term or critical access or transplant) for-profit hospital.

Source: CMS Provider of Service Files (December 2018)

Risk Factor	Mean	Minimum	Maximum
Indicator for presence of a for-profit hospital	.005	0	1

Number of hospitals: For each person, this variable records the number of active (short term or critical access or transplant) hospitals in the person's ZIP code tabulation area of residence.

Source: CMS Provider of Service Files (December 2018)

Risk Factor	Mean	Minimum	Maximum
Number of hospitals	.277	0	3

Number of hospitals per 1,000 residents: For each person, this variable records the number of active (short term or critical access or transplant) hospitals per 1,000 residents in the person's ZIP code tabulation area of residence.

Source: CMS Provider of Service Files (December 2018) American Community Survey (2017, 5-year estimates)

Risk Factor	Mean	Minimum	Maximum
Number of hospitals per 1,000 residents	.009	0	.4355401

Number of hospital beds per 1,000 residents: For each person, this variable records the number of active (short term or critical access or transplant) hospital beds per 1,000 residents in the person's ZIP code tabulation area of residence.

Source: CMS Provider of Service Files (December 2018) American Community Survey (2017, 5-year estimates)

Risk Factor	Mean	Minimum	Maximum
Number of hospital beds per 1,000 residents	1.894	0	65.80296

Number of primary care physicians per 1,000 residents: For each person, this variable records the number of primary care physicians per 1,000 residents in the person's ZIP code tabulation area of residence.

Source: AMA, American Community Survey (2017, 5-year estimates)

Risk Factor	Mean	Minimum	Maximum
Number of primary care physicians per 1,000 residents	.911	0	13.74227

Number of specialty care physicians per 1,000 residents: For each person, this variable records the number of specialty care physicians per 1,000 residents in the person's ZIP code tabulation area of residence.

Source: AMA, American Community Survey (2017, 5-year estimates)

Risk Factor	Mean	Minimum	Maximum
Number of specialty care physicians per 1,000 residents	1.834	0	44.88371

Maryland
DEPARTMENT OF HEALTH

The Hilltop Institute

Percentage aged 65 and over: For each person, this variable records the percentage of individuals in the person's ZIP code tabulation area of residence aged 65 and over (pooled from 2013-2017).

Source: American Community Survey (2017, 5-year estimates)

Risk Factor	Mean	Minimum	Maximum
Percentage aged 65 and over	15.582	0	85.71429

Percentage aged 0-4: For each person, this variable records the percentage of individuals in the person's ZIP code tabulation area of residence aged 0-4 (pooled from 2013-2017).

Source: American Community Survey (2017, 5-year estimates)

Risk Factor	Mean	Minimum	Maximum
Percentage aged 0-4	5.892	0	31

Percentage married: For each person, this variable records the percentage of the population aged 15+ in the person's ZIP code tabulation area of residence that is currently married (pooled from 2013-2017).

Source: American Community Survey (2017, 5-year estimates)

Risk Factor	Mean	Minimum	Maximum
Percentage married	48.079	0	91.4

Population growth: For each person, this variable records the percent population growth recorded in the person's ZIP code tabulation area of residence from 2011-2017.

Source: American Community Survey (2011 and 2017, 5-year estimates)

Risk Factor	Mean	Minimum	Maximum
Population growth	4.646	-100	286.1635

Population: For each person, this variable records the population of the person's ZIP code tabulation area of residence.

Source: American Community Survey (2017, 5-year estimates)

Risk Factor	Mean	Minimum	Maximum
Population	33350.99	0	106404

The Hilltop Institute

Population density: For each person, this variable records the population per square mile in the person's ZIP code tabulation area of residence.

Source: American Community Survey (2017, 5-year estimates), Census

Risk Factor	Mean	Minimum	Maximum
Population density	2842.714	0	95546.88

Percentage Native American: For each person, this variable records the percentage of the population in the person's ZIP code tabulation area that is Native American.

Source: American Community Survey (2017, 5-year estimates)

Risk Factor	Mean	Minimum	Maximum
Percentage Native American	.243	0	19.02389

Percentage non-white, aged 65+: For each person, this variable records the percentage of the population aged 65 and above in the person's ZIP code tabulation area that is non-white.

Source: American Community Survey (2017, 5-year estimates)

Risk Factor	Mean	Minimum	Maximum
Percentage non-white, aged 65+	27.888	0	98.98292

Percentage Hispanic, aged 65+: For each person, this variable records the percentage of the population aged 65 and above in the person's ZIP code tabulation area that is Hispanic.

Source: American Community Survey (2017, 5-year estimates)

Risk Factor	Mean	Minimum	Maximum
Percentage Hispanic, aged 65+	2.814	0	99.61651

Percentage with less than high school education, aged 65+: For each person, this variable records the percentage of the population aged 65 and above in the person's ZIP code tabulation area that has less than a high school diploma.

Source: American Community Survey (2017, 5-year estimates)

Risk Factor	Mean	Minimum	Maximum
Percentage with less than high school education, aged 65+	15.32	0	100

Percentage live alone, aged 65+: For each person, this variable records the percentage of the population aged 65 and above in the person's ZIP code tabulation area that lives alone.

Source: American Community Survey (2017, 5-year estimates)

Risk Factor	Mean	Minimum	Maximum
Percentage live alone, aged 65+	27.083	0	79.25

Percentage speak Spanish, aged 65+: For each person, this variable records the percentage of the population aged 65 and above in the person's ZIP code tabulation area that speaks Spanish.

Source: American Community Survey (2017, 5-year estimates)

Risk Factor	Mean	Minimum	Maximum
Percentage speak Spanish, aged 65+	2.608	0	95.69197

Percentage single mothers: For each person, this variable records the percentage of women aged 15-50 giving birth within the past 12 months who are not married in the person's ZIP code tabulation area.

Source: American Community Survey (2017, 5-year estimates)

Risk Factor	Mean	Minimum	Maximum
Percentage single mothers	32.952	0	100

Located in whole county primary care shortage area: For each person, this variable takes the value of 1 if the person's ZIP code tabulation area of residence is located in a county that is designated by HRSA in 2018 to be a whole-county primary care shortage area. The variable takes the value of 0, otherwise. If the ZIP code tabulation area lies in two or more counties, the value is estimated as a weighted average of the county-level attributes.

Source: Area Health Resources File

Risk Factor	Mean	Minimum	Maximum
Located in whole county primary care shortage area	0	0	1

Located in partial county primary care shortage area: For each person, this variable takes the value of 1 if the person's ZIP code tabulation area of residence is located in a county that is designated by HRSA in 2018 to be a partial-county primary care shortage area. The variable takes the value of 0, otherwise. If the ZIP code tabulation area lies in two or more counties, the value is estimated as a weighted average of the county-level attributes.

Source: Area Health Resources File

Risk Factor	Mean	Minimum	Maximum
Located in partial county primary care shortage area	.894	0	1.0001

Located in whole county mental health care shortage area: For each person, this variable takes the value of 1 if the person's ZIP code tabulation area of residence is located in a county that is designated by HRSA in 2018 to be a whole-county mental health care shortage area. The variable takes the value of 0, otherwise. If the ZIP code tabulation area lies in two or more counties, the value is estimated as a weighted average of the county-level attributes.

Source: Area Health Resources File

Risk Factor	Mean	Minimum	Maximum
Located in whole county mental health care shortage area	.198	0	1.0001

Located in partial county mental health care shortage area: For each person, this variable takes the value of 1 if the person's ZIP code tabulation area of residence is located in a county that is designated by HRSA in 2018 to be a partial-county mental health care shortage area. The variable takes the value of 0, otherwise. If the ZIP code tabulation area lies in two counties, the value is estimated as a weighted average of the county-level attributes.

Source: Area Health Resources File

Risk Factor	Mean	Minimum	Maximum
Located in partial county mental health care shortage area	.67	0	1.0001

General internists per 1,000 residents: For each person, this variable records the number of general internists per 1,000 residents in the county containing the ZCTA. If the ZIP code tabulation area lies in two or more counties, the value is estimated as a weighted average of the county-level attributes, with weights being the fraction of the ZCTA population residing within each county.

Source: Area Health Resources File

Risk Factor	Mean	Minimum	Maximum
General internists per 1,000 residents	.587	0	1.575016

The Hilltop Institute

Physician diversity: For each person, this variable records the percentage of medical doctors who are minorities (African Americans, Hispanics, and others, but excluding Asian-Americans). If the ZIP code tabulation area lies in two or more counties, the value is estimated as a weighted average of the county-level attributes, with weights being the fraction of the ZCTA population residing within each county.

Source: American Community Survey (2017, individual)

Risk Factor	Mean	Minimum	Maximum
Physician diversity	19.153	0	73.88705

Social workers per 1,000 residents: For each person, this variable records the number of social workers per 1,000 residents in the county containing the person's ZIP code tabulation area of residence. If the ZIP code tabulation area lies in two or more counties, the value is estimated as a weighted average of the county-level attributes, with weights being the fraction of the ZCTA population residing within each county.

Source: American Community Survey (2017, individual)

Risk Factor	Mean	Minimum	Maximum
Social workers per 1,000 residents	3.771	.5983358	6.583353

Appendix 2. HEDIS Antibiotics Tables

Source: HEDIS, Volume 2, pages 395-396.

Antibiotics of Concern by NCQA Drug Class Medications

Description	Prescriptions		
Quinolone	• Ciprofloxacin • Gemifloxacin	• Levofloxacin • Moxifloxacin	• Norfloxacin • Ofloxacin
Azithromycin and clarithromycin	• Azithromycin	• Clarithromycin	
Cephalosporin (second, third, fourth generation)	• Cefaclor • Cefdinir • Cefditoren • Cefepime • Cefixime	• Cefotaxime • Cefotetan • Cefoxitin • Cefpodoxime • Cefprozil	• Ceftriaxone • Cefuroxime • Ceftazidime • Ceftibuten
Amoxicillin/clavulanate	• Amoxicillin-clavulanate		
Ketolide	• Telithromycin		
Clindamycin	• Clindamycin		
Miscellaneous antibiotics of concern	• Aztreonam • Chloramphenicol	• Dalfopristin-quinupristin • Linezolid	• Telavancin • Vancomycin

All Other Antibiotics by NCQA Drug Class Medications

Description	Prescriptions	
Absorbable sulfonamide	• Sulfadiazine	• Sulfamethoxazole-trimethoprim
Aminoglycoside	• Amikacin • Gentamicin	• Streptomycin • Tobramycin
Cephalosporin (first generation)	• Cefadroxil • Cefazolin	• Cephalexin
Lincosamide (other than clindamycin)	• Lincomycin	
Macrolide (other than azithromycin and clarithromycin)	• Erythromycin • Erythromycin ethylsuccinate • Erythromycin lactobionate	• Erythromycin stearate • Erythromycin-sulfisoxazole
Penicillin (other than amoxicillin/clavulanate)	• Ampicillin • Ampicillin-sulbactam • Amoxicillin • Dicloxacillin • Nafcillin	• Penicillin G potassium • Penicillin G procaine • Penicillin G sodium • Penicillin V potassium • Piperacillin-tazobactam

The Hilltop Institute

	• Oxacillin • Penicillin G benzathine	• Ticarcillin-clavulanate
Tetracyclines	• Doxycycline • Minocycline	• Tetracycline
Miscellaneous antibiotics	• Daptomycin • Fosfomycin • Metronidazole • Nitrofurantoin	• Nitrofurantoin macrocrystals • Rifampin • Trimethoprim

The Hilltop Institute

UMBC

Sondheim Hall, 3rd Floor
1000 Hilltop Circle
Baltimore, MD 21250
410-455-6854
www.hilltopinstitute.org