Informing Medicare with Evidence on Social Risk Adjustment

Technical Specifications

Prepared for:

The Patrick and Catherine Weldon Donaghue Medical Research Foundation

18 North Main Street

West Hartford, Connecticut 06107

Prepared by:

Danielle Browne, BS, Kano Amagai, MPH, Austin Kilaru, MSHP, MD,
Aidan Crowley, BS, Amol Navathe, MD, PhD, Joshua Liao, MD, MSc

July 2025

Purpose

This technical specification document outlines the proposed objectives and methodology of research work supported by the Patrick and Catherine Weldon Donaghue Medical Research Foundation. The specifications below will guide manuscripts, conference presentation, and issue brief deliverables.

Background and Objectives

Study Context

The goal of this research is to fill critical knowledge gaps in value-based care and health equity with evidence about social risk adjustment, i.e., risk adjustment that accounts for social needs that may contribute to health outcomes. As the largest payer in the US, the Centers for Medicare & Medicaid Services (CMS) is leading such efforts by incorporating social risk adjustment into large nationwide payment programs using the Area Deprivation Index (ADI) – a composite, neighborhood-level measure of socioeconomic disadvantage that encompasses income, employment, education, housing, transportation, and other social determinants of health. Our work aims to fill Medicare-specific knowledge gaps on social risk adjustment to promote policy and practice change with empirical evidence.

Analytic Objectives

First, we will examine whether there is an association between neighborhood socioeconomic disadvantage (higher ADI) and patient (mortality, hospitalization) and cost outcomes (total spending) despite adjustment for clinical factors, establishing whether there is need to augment traditional clinical risk adjustment strategies.

As part of this work, we will also evaluate recent modifications of the ADI, including the construction of a related index that improves the underlying methodology, including standardization of the underlying component variables measuring neighborhood socioeconomic disadvantage.

Data Sources

Primary Datasets

The following sources are to be used: Medicare Fee-for-Service claims (inpatient, outpatient, carrier, skilled nursing facility), Medicare Beneficiary Summary File; Medicare Data on Provider Practice and Subspecialty; 2023 Medicare Enrollment DataBase (EDB) file; American Community Survey 2020 (US Census Bureau); 2020 Neighborhood Atlas version 4.0.1. (University of Wisconsin). The population and respective timeframe will include a 20% random sample of Medicare Fee-for-Service beneficiaries continuously enrolled in Medicare Part A and Part B from January 2018 through January 2019.

Data Linkage

The methodology to complete the proposed work will entail both combining and linking of various datasets and variables. Procedures may include the following:

- Linking of physician-level National Provider Identification associated with outpatient claims to the Medicare Data on Provider Practice and Subspecialty data to describe outpatient healthcare utilization

- Linking 9 digit-zip level ADI to the 9-digit zip-code of beneficiaries' residence listed on the EDB file

- Constructing standardized ADI for each census block group using American Community Survey data and use of the 9-digit zip to Census Block Group crosswalk in the Neighborhood Atlas to link the standardized ADI to each beneficiary based on their 9-digit zip in the EDB file

- Rural-Urban Commuting Area codes, based on 5-digit zip-code, to be linked to the first 5 digits of the beneficiaries' residence listed on the EDB file

Data Access & Permissions

Data permissions are granted through a Data Use Agreement between the University of Pennsylvania and the CMS Research Data Assistance Center (ResDAC). Data is accessible through ResDAC's Chronic Conditions Warehouse Virtual Research Data Center (VRDC). The VRDC satisfies all CMS privacy and security requirements and allows contracted users to access Research Identifiable Files through a secure cloud. To ensure the privacy and confidentiality of all project data accessed outside of the VRDC (e.g., ADI information), we store, use, and analyze identifiable data only on a secure server located in the Health Services Research Data Center, a high-security server center shared by the University of Pennsylvania School of Medicine and the University of Pennsylvania Health System. For more information, please visit: https://resdac.org/cms-virtual-research-data-center-vrdc.

Cohort Construction

Inclusion and Exclusion Criteria

Inclusion criteria are Medicare Fee-for-Service beneficiaries continuously enrolled in Medicare Part A and Part B from January 2018 through January 2019.

Exclusion criteria are Medicare beneficiaries enrolled in Medicare Advantage; beneficiaries with end-stage-renal-disease or age greater than 100 years; beneficiaries who could not be distinguished as either fully or partially dual-eligible for Medicare and Medicaid or who qualified for Medicare only due to disability status; beneficiary claims with non-Medicare primary payer; beneficiaries missing geographic information.

For the spending outcomes cohort, in addition to the exclusion criteria listed above, we will remove Maryland residents (due to a state-specific payment model) and those with unavailable standardized cost information.

Index Dates

Outcomes are to be measured starting from January 2019. All baseline information is based on information collected in 2018. Observation time ends when the beneficiary leaves Medicare Fee-for-Service enrollment or the end of 2019 (whichever occurs first).

Key Variables and Measures

Exposure(s)

The primary exposure will include the Area Deprivation Index (ADI) value assigned to each beneficiary based on the 9-digit zip code of their residence listed in the Enrollment Data Base prior to January 2019.

Outcome(s)

The primary outcome includes all-cause mortality in 2019, using the Medicare Beneficiary Summary File to identify beneficiary death dates. Secondary outcomes include hospital admissions and total spending in 2019 using claims for inpatient, outpatient, physician, post-acute care, home health, hospice, and durable medical equipment services. For total spending, we will adjust for the duration of individual Medicare Fee-for-Service coverage during 2019 to derive total spending per beneficiary per month. All spending outcomes are to be standardized to account for regional differences.

Covariates

Planned confounders and their operational definitions are listed below.

- Age: report continuously in years as measured in 2018

- Sex: Male or Female

- US region of residence: Northeast, South, Midwest, or West

- Rural-urban community area codes: Metropolitan, Micropolitan, Small town, Rural

- Dual eligibility for Medicaid and Medicare in 2018: partial dual, full dual, or non-dual

- Original reason for Medicare entitlement: old age, disability

- CMS' Hierarchical Condition Categories (CMS-HCC) v24: 86 variables that group together ICD-10 diagnosis codes that share similar characteristics in terms of predicting healthcare costs

Variable Construction Logic

CMS-HCC to be constructed using inpatient, outpatient, carrier and skilled nursing facility claim files from 2019 and created using the logic listed in this website: https://www.cms.gov/medicare/payment/medicare-advantage-rates-statistics/risk-adjustment. Construction of standardized ADI using downloads of 5-year American Community Survey data from 2020 for Census block group, Census tract, and county. We will follow methodologies published in this article: https://academic.oup.com/healthaffairsscholar/article/1/5/qxad063/7342005.

Analytic Approach

Study Design

The proposed work entails a cross-sectional analysis to compare outcomes among Medicare beneficiaries across distributions of ADI.

Primary Analysis Methods

A beneficiary-level ordinary least squares regression model will be used to estimate mortality in 2019 and determine the relationship between ADI and excess mortality. The model will include CMS-HCC and ADI distribution.

Handling of Missing Data

Medicare beneficiaries missing geographic information will be excluded from our analysis as noted under exclusion criteria.

Sensitivity/Secondary Analyses

For robustness, there are two secondary analyses planned. First, a beneficiary-level ordinary least squares regression model may be used to estimate hospital admission in 2019 and determine the relationship between ADI and excess hospitalization. The model may include CMS-HCC and ADI distribution. Second, a beneficiary-level ordinary least squares regression model may be used to estimate standardized spending per beneficiary per month in 2019 and determine the relationship between ADI and excess spending. The model may include CMS-HCC and ADI distribution.

Sensitivity analyses may include the following procedures:

- Repeat primary analysis using standardized ADI

- Stratify main analysis by dual-eligibility status

- Stratify main analysis by rurality (separate beneficiaries located in metropolitan 5-digit zip codes from those living in non-metropolitan areas)

- Repeat main analysis using the updated v28 of the CMS-HCC

- Repeat main analysis using hospital-free days as an additional secondary outcome using a Tobit regression model

Statistical Tools and Software

Analyses are to be performed with SAS version 9.4.

Dissemination

Output and Deliverables

Planned outputs for the proposed bodies of work (e.g., manuscript(s), issue brief(s), conference presentation(s)) may include tables and figures that capture the following characteristics or information:

- Patient characteristics, overall and by ADI distribution

- Unadjusted and/or adjusted mortality, hospital admissions, and per beneficiary per month spending, overall and by ADI distribution

- Adjusted excess mortality overall and by ADI distribution, adjusted for HCC score

- Adjusted excess hospital admissions overall and by ADI distribution, adjusted for HCC score

- Adjusted excess total spending overall and by ADI distribution, adjusted for HCC score

Public Sharing Plans

Platforms for dissemination may include the Leonard Davis Institute of Health Economics; University of Pennsylvania and University of Texas Southwestern Medical Center digital platforms. Additional dissemination activities may include real-time discussions, webinars, and/or policy briefings to disseminate findings with relevant stakeholders.