

# Technical Recommendations for Advancing Health Equity with PQA Quality Measures:

A Report from the PQA Health Equity Technical Expert Panel



---

Ben Shirley, CPHQ  
Melissa Castora-Binkley, PhD  
Carly Vogel, PhD  
Razanne Oueini, PharmD, MSc, CPHQ  
Lisa Hines, PharmD, CPHQ



# Table of Contents

|   |    |
|---|----|
| <b>Executive Summary</b>  | 3  |
| <b>Introduction</b>   | 6  |
| <b>Purpose and Objectives</b>   | 7  |
| <b>TEP Recruitment, Structure, and Process</b>  | 9  |
| <b>Objective 1   Stratification in PQA Measure Testing</b>  | 10 |
| PQA Measure Lifecycle and Testing   | 10 |
| Stratification Basics   | 10 |
| TEP Discussions and Recommendations   | 11 |
| Next Steps  | 15 |
| <b>Objective 2   Stratification of PQA Measures in Quality Programs</b>                                     | 15 |
| Medicare Part D Star Ratings  | 15 |
| Medicaid Adult Core Set   | 17 |
| Health Insurance Marketplace Quality Rating System  | 18 |
| Integrated Healthcare Association Align. Measure. Perform. Program  | 19 |
| Next Steps  | 19 |
| <b>Objective 3   Affirming Existing PQA Adherence Risk Adjustment Model and Exploring Future Directions</b> | 20 |
| Risk Adjustment Overview  | 20 |
| History of PQA Risk Adjustment of Adherence Measures  | 20 |
| TEP Discussions and Recommendations   | 22 |
| Next Steps  | 22 |
| <b>Objective 4   Feasibility of Risk Adjusting the PDC-CMP Health Plan Measure</b>                          | 23 |
| Background  | 23 |
| Subject Matter Expert and TEP Discussions and Recommendations   | 23 |
| Next Steps  | 24 |
| <b>Conclusions and Considerations for Future Work</b>   | 25 |
| <b>About PQA</b>  | 26 |
| <b>References</b>   | 27 |
| <b>Appendix: TEP Discussion Details on Stratification Variables for PQA Measure Testing</b>                 | 30 |

# Executive Summary

Health equity has grown as a critical area of focus in the United States healthcare system. Numerous groups have been convened to advance health equity in healthcare quality measurement. However, limited information is available that is specific and immediately actionable for PQA measures. To meet this need, and to advance its health equity strategy and integrate an equity focus into measurement related activities, the PQA Health Equity Technical Expert Panel (TEP) was convened to focus on practical and meaningful technical approaches.

Appointed TEP members included representatives from both PQA member and nonmember organizations. Collectively, TEP members provided diverse perspectives from a wide variety of stakeholders, including health plans, pharmacies, life sciences organizations, pharmacy benefit managers, health systems, professional associations, universities, and health technology organizations, among others. Importantly, the TEP also included patient representatives. TEP webinar meetings were held between June 2023 and March 2024 and focused on four key objectives:

**I. Identify stratification variables for standard inclusion in PQA's measure testing process**

**II. Evaluate and make recommendations regarding the use of identified stratification approaches for PQA measures in quality programs**

**III. Review and affirm the risk adjustment model currently recommended for use with the Proportion of Days Covered health plan measures (Statins, Diabetes All Class, Renin-Angiotensin System [RAS] Antagonists) in Medicare Part D**

**IV. Explore the feasibility of applying a risk adjustment model to the *Proportion of Days Covered: Composite* (PDC-CMP) health plan measure**

## **I. Identify stratification variables for standard inclusion in PQA's measure testing process**

PQA measures are comprehensively tested prior to endorsement and use in quality programs. By stratifying testing results by the appropriate sociodemographic variables, PQA can identify potential disparities and promote measure use to improve care among vulnerable subgroups. The TEP evaluated a variety of sociodemographic variables to identify those that should be included as standard within PQA's testing process. Discussions focused on whether a variable was feasible to be used for consistent stratification, and how important a variable was to stratify in terms of impact on the measure. The TEP also discussed stratification structures for these variables at length.

**Age, sex, and race and ethnicity** were consistently prioritized by the TEP, while primary language was least prioritized, and geography, income, disability, and dual eligibility for Medicare and Medicaid ranked as medium priorities.

## II. Evaluate and make recommendations regarding the use of identified stratification approaches for PQA measures in quality programs

While identifying disparities in PQA's testing processes is important, it is not sufficient. Quality measures are most impactful when used in quality programs that drive interventions and improvement. Stratifying by sociodemographic variables in these programs can provide valuable information to patients while also identifying disparities within and across measured entities.

The TEP made program-specific recommendations for stratified reporting of PQA measures used in quality programs, including the **Medicare Part D Star Ratings, Medicaid Adult Core Set, Health Insurance Marketplace Quality Rating System (QRS), and Integrated Healthcare Association (IHA) Align. Measure. Perform. (AMP) program**. As with the previous objective, discussions centered on variables that are both important and feasible, and recommendations were provided in the form of a prioritized list of variables for stratification.

The TEP's recommendations were consistent across programs, with **age, sex, and geography** most highly prioritized for the Medicare Part D Star Ratings, the Medicaid Adult Core Set, Health Insurance Marketplace QRS, and IHA AMP programs. For most programs, **race and ethnicity** was the next highest priority.

## III. Review and affirm the risk adjustment model currently recommended for use with the Proportion of Days Covered health plan measures (Statins, Diabetes All Class, Renin-Angiotensin System [RAS] Antagonists) in Medicare Part D

PQA previously developed a sociodemographic risk adjustment model for the three PQA Proportion of Days Covered (adherence) measures used in the Medicare Part D Star Ratings program (Statins, Diabetes All Class, RAS Antagonists). In a proactive response to the planned implementation of this risk adjustment model in Medicare Part D Star Ratings program for 2028, the TEP reviewed and affirmed the existing risk adjustment model used for all PQA Proportion of Days Covered measures.

Collectively, **the TEP reviewed the current risk adjustment model and affirmed its continued importance and appropriateness**, noting that its implementation in the Part D Star Ratings should be recognized as a success. For future iterations of the risk adjustment model that may be developed in the future, **race and ethnicity** was the most recommended variable, although several other variables were also discussed.

## IV. Explore the feasibility of applying a risk adjustment model to the Proportion of Days Covered: Composite (PDC-CMP) health plan measure

PQA developed the PDC-CMP measure, combining rates for three PQA adherence measures (Statins, Diabetes All Class, RAS Antagonists) in response to stakeholder support for a summary indicator of these adherence measures. However, the risk adjustment model developed for individual adherence measures was not designed for use with the PDC-CMP measure. The TEP, in

conjunction with external statistical subject matter experts, evaluated the feasibility of applying a risk adjustment model to the PDC-CMP measure.

The TEP concluded that **developing a risk adjustment model for the PDC-CMP measure could be feasible** and discussed conceptual considerations for risk adjusting the PDC-CMP measure in comparison to individual adherence measures. The TEP ultimately recommended using the variables currently used in the existing adherence risk adjustment model, and additional variables like **race and ethnicity**, as a starting point for the PDC-CMP measure. However, given the required resources, future development of a risk adjustment model for the PDC-CMP measure would likely be contingent on use of the measure in a national quality program.

Overall, the TEP provided valuable health equity recommendations for PQA measure testing and use in quality programs, while also providing useful direction for future iterations of risk adjustment models. Health plans and quality programs can apply these recommendations to identify disparities in their performance and build or incentivize interventions to enhance medication use quality and equity. Other users of PQA measures, such as researchers and educators, can also use these recommendations in their work, while quality measure developers can use these recommendations as a resource when exploring measure stratification.

Finally, PQA recognizes and emphasizes that identification of disparities is not enough to achieve health equity. Quality measurement is a critical tool to highlight existing disparities and track progress, but it must be accompanied by investments and interventions to address gaps and improve care provided to vulnerable subgroups.

# Introduction

Health equity has grown as a critical area of focus in the United States healthcare system. Executive order 13985 established a federal priority to pursue a comprehensive approach to advance equity for all, specifically calling out health equity and aiming to expand quality health care to vulnerable populations. In turn, the Department of Health and Human Services (HHS) reinforced its emphasis on the advancement of health equity through the 2023 HHS Equity Action Plan, which called on the healthcare system to “acknowledge and incorporate equity considerations into what we do in order to optimize opportunities for healthy and thriving lives.”<sup>1</sup> Given that health equity is a dimension of care quality within itself, it is essential to consider how quality measures can be further leveraged to promote health equity for all individuals. **PQA committed to addressing health equity in its current and future measurement activities as a pillar of its strategic plan, Blueprint PQA 2025.**

A number of reports have been published in the last decade aiming to provide guidance on how to address issues of disparities and health equity in quality measures through approaches such as risk adjustment and stratification.<sup>2-6</sup> In 2014, the National Quality Forum (NQF) recommended that each quality measure be assessed individually to determine the appropriateness of risk adjustment and also recommended that measures include specifications for stratification. In 2016, the Office of the Assistant Secretary for Planning and Evaluation (ASPE) published a report that stated the need for enhanced data collection, health equity measures, and more research to determine the appropriateness of quality measure risk adjustment. ASPE followed with a second report in 2020 that stated measures should not be risk adjusted for social factors in public reporting or value-based programs.

Contrary to ASPE’s most recent position, in 2021, NQF established minimum requirements to evaluate specific social risk factors during measure development, focusing on the importance of a conceptual model.<sup>2</sup> A conceptual model illustrates the pathways and connections between social risk factors, clinical risk factors, clinical process of interest, and an associated outcome. Social risk factors of interest should be included based on a logical conceptual rationale, not simply a significant statistical test result.

This process of evaluating social risk factors during measure development should also delineate which potential mechanisms affecting the outcome are actionable by the measured entity, versus those that are intrinsic patient factors present at the start of care that influence the health of the patient and their outcomes. This latter category of factors, which falls outside the direct control of measured entities, has historically faced a lack of consensus across the industry regarding its role in quality measurement.

## KEY PQA HEALTH EQUITY TEP RECOMMENDATIONS

### *for PQA measure testing and use in quality programs*

- **Age, sex, and race and ethnicity** are the top recommended variables for stratifying PQA measure testing results, based on feasibility and importance.
- **Age, sex, and geography** are the most highly recommended variables for stratified reporting of PQA measures used in quality programs, based on immediate feasibility and readiness.

### *for PQA risk adjustment models*

- The current sociodemographic risk adjustment model for PQA adherence measures is meaningful and impactful, and future iterations should consider **race and ethnicity** among other variables, pending data quality and availability.
- A risk adjustment model for the **PDC-CMP** measure could be feasible, and the existing adherence risk adjustment model variables and **race and ethnicity** should be prioritized.



Nevertheless, NQF stated that developers at a minimum should consider certain risk factors for inclusion during the development of the conceptual model, including age, gender, urbanicity/rurality, poverty, social vulnerability indices, and indicators of frailty and disability. Race and ethnicity were the subject of considerable debate, although NQF ultimately recommended considering race and ethnicity for inclusion in conceptual model development as well. Further, NQF stated that quality measures should be stratified in conjunction with risk adjustment.

Risk adjusting quality measures for social risk factors can be considered more impactful for measured entities compared to patients, as it aims to allow fairer comparisons between measured entities. PQA has provided recommendations for risk adjustment of three PQA Proportion of Days Covered (adherence) measures used in the Medicare Part D Star Ratings program (Statins, Diabetes All Class, RAS Antagonists) and these recommendations have been adopted by CMS. Stratification, on the other hand, is often considered a more patient-focused endeavor as it helps to identify subgroups where differences in performance exist.<sup>2</sup> These differences often represent disparities in care, which can be targeted for interventions to improve processes and outcomes. The use of measures to identify disparities through stratification and addressing with subsequent targeted quality improvement is a key strategy to promote health equity for patients.

Given the continued importance of advancing health equity and existing recommendations for measure developers to provide guidance for stratification, this report summarizes the results of the PQA Health Equity Technical Expert Panel (TEP) to provide specific health equity recommendations for PQA measures. Similar to cited guidance, these recommendations are not meant to be prescriptive and are not exhaustive of approaches to achieve health equity. PQA recognizes the variability in data availability and quality, but notwithstanding these challenges, the TEP recommendations provide PQA with immediate and actionable steps to meet the organization's commitment to health equity.

The content of this report is meaningful to a variety of audiences. This report will help guide users of PQA measures, such as CMS and health plans, to investigate where disparities may exist. Organizations focused on quality improvement can use this guidance to identify the existence of disparities and make recommendations to enhance quality so that all individuals are able to optimize their medication use. Researchers can use these recommendations in studies to shine a light on performance within key subgroups, and educators can use the information to broaden awareness of the complexities, challenges, and opportunities to enhance health equity through quality measurement.

## Purpose and Objectives

---

Several groups have convened to advance health equity in healthcare quality measurement. However, limited information is available that is specific and immediately actionable for PQA measures. The PQA-convened Health Equity TEP was designed to focus on practical and meaningful technical approaches that advance PQA's health equity strategy and integrate an equity focus into existing activities. Necessarily, the TEP prioritized a limited number of discussion topics based on where their input and recommendations could be considered for rapid implementation.

## **OBJECTIVE 1 | Identify stratification variables for standard inclusion in PQA's measure testing process.**

PQA measures are comprehensively tested prior to endorsement and use in quality programs. By stratifying testing results by the appropriate sociodemographic variables, PQA can identify potential disparities and promote measure use to improve care among vulnerable subgroups. The TEP was tasked with evaluating sociodemographic variables for standard inclusion in PQA's testing process. The TEP was also tasked with recommending a stratification structure to operationalize these variables (i.e., establishing consistent variable definitions).

## **OBJECTIVE 2 | Evaluate and make recommendations regarding the use of identified stratification approaches for PQA measures in quality programs.**

The identification of disparities in the PQA testing processes is important; however, it is not sufficient. Quality measures are most impactful when used in quality programs that drive interventions and improvement. Sociodemographic variable stratification of measures used in these programs can provide valuable information to patients while highlighting disparities within and across measured entities. The TEP was tasked with evaluating sociodemographic variables for major quality programs in which PQA measures are used and making program-specific recommendations for stratified reporting of PQA measures.

## **OBJECTIVE 3 | Review and affirm the risk adjustment model currently recommended for use with the Proportion of Days Covered health plan measures (Statins, Diabetes All Class, Renin-Angiotensin System [RAS] Antagonists) in Medicare Part D.**

PQA previously developed a sociodemographic risk adjustment model for the three PQA Proportion of Days Covered (adherence) measures used in the Medicare Part D Star Ratings program (Statins, Diabetes All Class, RAS Antagonists). The risk-adjusted versions of these adherence measures will be implemented in the Medicare Part D Star Ratings program in Star Ratings year 2028.<sup>7</sup> The TEP was tasked with revisiting the original recommendations for variables included in the risk adjustment model, reaffirming their continued importance, and evaluating new variables for inclusion in potential future iterations of the risk adjustment model.

## **OBJECTIVE 4 | Explore the feasibility of applying a risk adjustment model to the Proportion of Days Covered: Composite (PDC-CMP) health plan measure.**

PQA developed a composite PDC health plan measure that combines rates for three PQA adherence measures (Statins, Diabetes All Class, RAS Antagonists) in response to stakeholder support for a summary indicator of these three PDC measures. As noted above, PQA developed a risk adjustment model for the individual adherence measures included in the PDC-CMP measure. However, the development and application of a risk adjustment model for composite measures is complex. The TEP, in conjunction with external statistical subject matter experts (SMEs), was tasked with evaluating the feasibility of applying a risk adjustment model to the PDC-CMP measure.



# TEP Recruitment, Structure, and Process

PQA used a multi-faceted TEP recruitment strategy, including a call for self-nominations (March 20 to April 7, 2023) distributed via multiple communications channels, and targeted outreach inviting SMEs with relevant expertise to self-nominate. PQA specifically sought individuals with experience in promoting health equity, quality measurement, statistical methodologies, and serving patients from diverse backgrounds. PQA's [Principles for Diversity, Equity and Inclusion](#) were applied during the selection process, as PQA aims to appoint panels that represent the nation's diversity and are responsive to the needs and interests of diverse populations.

PQA staff selected 21 individuals to serve on the Health Equity TEP. Appointed TEP members included representatives from both PQA member organizations and nonmember organizations and provided diverse perspectives from a wide variety of stakeholders, including health plans, pharmacies, pharmaceutical industry organizations, pharmacy benefit managers, health systems, professional associations, universities, and health technology organizations, among others. Importantly, the TEP also included patient representatives.

Six TEP meetings were held via webinar between June 2023 and March 2024. PQA staff, in collaboration with chairs, collated and presented relevant background information to the TEP. Each meeting focused on one of the four TEP objectives, with some objectives requiring multiple meetings for discussion. TEP members were asked to complete post-meeting surveys that formalized their recommendations and considerations for each objective. The TEP roster is provided below, with individuals attributed to the organizations they were associated with at the time of the TEP's convening.

- Sujith Ramachandran, *University of Mississippi (co-chair)*
- Christie Teigland, *Inovalon (co-chair)*
- Elle Blouin, *Aetna*
- Matthew Dinh, *SCAN Health Plan*
- Marybeth Farquhar, *American Urological Association*
- Dana Erf Fortman, *Walgreens*
- Denean Green Rivers, *Patient Partner*
- Katie Herndon, *Pfizer, Inc.*
- Michelle Juhanson, *Magellan Rx Management*
- Shellie Keast, *American Drug Utilization Review Society*
- Aimee Loucks, *Kaiser Permanente*
- Nate Lucena, *Rex Wallace Consulting*
- Erin Neal, *Vanderbilt University Medical Center*
- Dani Markus, *Cardinal Health/Outcomes*
- David Parker, *Johnson & Johnson*
- Kamal Patel, *Elevance Health*
- Niki Shah, *McKesson*
- Kyle Thompson, *Prime Therapeutics LLC*
- Andrew Thorne, *Pharmacy Quality Solutions*
- George Valentine, *Patient Partner*
- Courtney Walker, *NovoNordisk, Inc.*

# Objective 1 |

## Stratification in PQA Measure Testing

---

### PQA Measure Lifecycle and Testing

To contextualize the TEP's recommendations for stratification within measure testing, it is important to understand how testing fits into the PQA measure lifecycle. PQA's measure lifecycle consists of five non-linear phases: conceptualization, specification, testing, endorsement, and maintenance and use. While measures typically move through different phases as they advance through development, the lifecycle is iterative, and measures may move back and forth between phases.

Throughout the measure lifecycle, PQA focuses on four key measure evaluation criteria. First, the *importance* of a measure is evaluated to ensure it is focused on a priority area, rooted in the evidence, and can have a positive impact on healthcare quality. Second, the measure must be *scientifically acceptable* and produce reliable and valid results about the quality of care. Third, the *feasibility* of the measure is considered, including calculation burden and whether data are readily available for measurement and retrievable without undue burden. Finally, the *usability* of a measure is assessed to ensure there is opportunity for implementation and performance results can be used for both accountability and performance improvement.

To address scientific acceptability during PQA's measure development process, measure concepts undergo **measure testing**. In this phase, PQA develops and executes a comprehensive measure testing plan to evaluate reliability, validity, feasibility, and performance gaps. PQA regularly partners with external organizations as volunteers to test the measure concept, including health plans, pharmacy benefit managers, vendors/data aggregators, and pharmacies (for pharmacy measurement). PQA also uses internal data assets, such as Medicare or Medicaid samples. Testing typically takes place across three lines of business: Medicare, Medicaid, and commercial. Beyond development, measures for which substantive changes are under consideration as part of maintenance may also undergo testing.

### Stratification Basics

Stratification refers to the arrangement of performance scores for specific population subgroups. It provides greater attention to potential disparities and enables targeted quality improvement and accountability. Generally, no statistical modeling is required for stratification. For example, in addition to reporting one overall performance score, an accountable entity would include multiple performance scores, one for each stratum. Stratifications are often related to social and functional status variables and are often used to highlight differences in care by comparing scores across strata. Stratification of differences among subgroups within a single entity can promote health equity by identifying opportunities for appropriate resource allocation for quality improvement and for reducing disparities in care delivery.

## TEP Discussions and Recommendations

The first objective of the TEP was to recommend variables for standard inclusion in PQA measure testing, with an emphasis on evaluating appropriate stratification structures. Discussion focused primarily on age, sex, race and ethnicity, geography, dual-eligible status, an income indicator, disability, and primary language. Some of these variables are already included in PQA measure testing in various structures, while others would be entirely new.

Following discussions, the TEP engaged in a survey (86% response rate) to prioritize the recommended stratification variables by their feasibility (e.g., ability to be consistently captured and used) and importance (impact on health equity). Because measure testing is resource-intensive for both PQA staff and external partners, prioritization is critical to focus PQA’s efforts on adding and refining the most important and feasible variables to measure testing processes.

### Variable Prioritization and Recommended Stratification Structures

Tables 1 and 2 provide the full results of the ranking exercise.

| TABLE 1   Prioritization of Recommended PQA Testing Stratification Variables by Feasibility |                  |                |              |                 |                 |  |   |
|---|------------------|----------------|--------------|-----------------|-----------------|--|---|
| PRIORITY LEVEL  | FEASIBILITY      | MEDIAN RANKING | MEAN RANKING | MINIMUM RANKING | MAXIMUM RANKING | PERCENTAGE RESPONDENTS RANKED IN TOP 3 | PERCENTAGE RESPONDENTS RANKED IN BOTTOM 3 |
| HIGHEST   | AGE              | 1.0            | 2.2          | 1               | 7               | 83.3%                                  | 11.1%                                     |
|   | SEX              | 2.0            | 2.1          | 1               | 3               | 100.0%                                 | 0.0%                                      |
|   | RACE/ETHNICITY   | 3.5            | 3.8          | 1               | 8               | 50.0%                                  | 22.2%                                     |
|   | DUAL ELIGIBILITY | 4.0            | 4.6          | 1               | 8               | 38.9%                                  | 38.9%                                     |
|   | GEOGRAPHY        | 4.5            | 4.6          | 1               | 6               | 11.1%                                  | 27.8%                                     |
|   | INCOME           | 5.5            | 5.5          | 2               | 8               | 11.1%                                  | 50.0%                                     |
|   | DISABILITY       | 6.5            | 6.3          | 3               | 8               | 5.6%                                   | 61.1%                                     |
| LOWEST  | PRIMARY LANGUAGE | 7.0            | 6.9          | 4               | 8               | 0.0%                                   | 88.9%                                     |

.....

| TABLE 2   Prioritization of Recommended PQA Testing Stratification Variables by Importance |                  |                |              |                 |                 |  |   |
|--|------------------|----------------|--------------|-----------------|-----------------|--|---|
| PRIORITY LEVEL   | IMPORTANCE       | MEDIAN RANKING | MEAN RANKING | MINIMUM RANKING | MAXIMUM RANKING | PERCENTAGE RESPONDENTS RANKED IN TOP 3 | PERCENTAGE RESPONDENTS RANKED IN BOTTOM 3 |
| HIGHEST  | RACE/ETHNICITY   | 2.0            | 2.3          | 1               | 5               | 77.8%                                  | 0.0%                                      |
|  | AGE              | 3.0            | 3.5          | 1               | 7               | 55.6%                                  | 38.9%                                     |
|  | SEX              | 4.0            | 4.1          | 1               | 8               | 38.9%                                  | 22.2%                                     |
|  | INCOME           | 4.0            | 3.9          | 1               | 8               | 44.4%                                  | 16.7%                                     |
|  | GEOGRAPHY        | 4.0            | 4.4          | 1               | 8               | 27.8%                                  | 27.8%                                     |
|  | DISABILITY       | 6.0            | 5.8          | 3               | 8               | 11.1%                                  | 55.6%                                     |
|  | DUAL ELIGIBILITY | 6.0            | 5.6          | 2               | 8               | 27.8%                                  | 61.1%                                     |
| LOWEST   | PRIMARY LANGUAGE | 7.0            | 6.4          | 2               | 8               | 16.7%                                  | 77.8%                                     |

**Age, sex, and race and ethnicity** were ranked in the top three for both feasibility and importance. Note that although race and ethnicity are distinct, they are often referred to together (i.e., *race* and *ethnicity*) due to their close conceptual relationship. For ease of reading, they are referred to in tandem in this report except when content is specific to only one of the two concepts.

## Recommended Stratification Structures

### AGE

- **Current stratification structure in PQA testing:** 18-50, 51-64, 65-84, 85+
- **Recommended stratification structure:** 18-34, 35-49, 50-64, 65-74, 75-84, and 85+
  - The TEP considered this stratification structure a starting point and encouraged tailoring it by measure as appropriate based on clinical factors or measure focus.

### SEX

- **Current stratification structure:** Male, Female
- **Recommended stratification structure:** Male, Female
  - The TEP agreed this variable was currently most feasible to stratify as male or female, although it is less meaningful without including an “other,” “unknown,” or “prefer not to disclose” option.
  - The TEP recommended addition of more granular sexual orientation or gender identity (SOGI) information should it become more feasible in the future.

### RACE

- **Current stratification structure:** Not included in standard PQA testing
- **Recommended stratification structure:** American Indian or Alaska Native, Asian, Black or African American, Native Hawaiian or Other Pacific Islander, White

### ETHNICITY

- **Current stratification structure:** Not included in standard PQA testing
- **Recommended stratification structure:** Hispanic or Latino, Not Hispanic or Latino
  - The TEP expressed support for using Office of Management and Budget (OMB) categories as a starting point for race and ethnicity, but suggested re-evaluating as data improves.
  - The TEP recommended future considerations to include “other” as an option, the inclusion of a category for multi-racial, mixed race, multiple races, and the use of more granular categories for Asian.
  - As noted in the discussion details, the OMB categories have been updated since the TEP discussions. PQA plans to align with the updated OMB categories.

**Geography** was consistently rated as moderate feasibility and importance. **Income** was ranked in the bottom three for feasibility but higher for importance, and **dual eligibility** was ranked in the bottom three for importance but higher in feasibility.

## Recommended Stratification Structures

### GEOGRAPHY

- **Current stratification structure:** Not included in standard PQA testing
- **Recommended stratification structure:** No standard structure identified
  - While the TEP discussion centered strongly on urban, suburban, and rural designations, no consensus was reached on how to define these classifications. Numerous definitions and methodologies exist to evaluate urbanicity and rurality.
  - TEP members noted that while using ZIP code mapping to assign patients to rural, urban, or suburban values is possible, it would be burdensome to testers.
  - Consensus was not reached on a standard stratification structure that reflects consistent definitions of rural, urban, or suburban.

### DUAL ELIGIBILITY

- **Current stratification structure:** Dual Eligible, Non-Dual Eligible (Medicare and Medicaid only)
- **Recommended stratification structure:** Dual Eligible, Non-Dual Eligible (Medicare and Medicaid only)
  - The TEP ranked dual eligibility's importance relatively lower due to its largely duplicative nature with income. However, the TEP noted that the variable is feasible and impactful in Medicare and Medicaid populations.

### INCOME

- **Current stratification structure:** Low-Income Subsidy (LIS), Non-LIS (Medicare only)
- **Recommended stratification structure:** Low-Income Subsidy (LIS), Non-LIS (Medicare only)
  - The TEP agreed on the feasibility and ease of capturing LIS or non-LIS as an income indicator within Medicare. The TEP commented that capturing income more precisely with a dollar amount or range is more meaningful and important, but less feasible.
  - Income data outside Medicare is less feasible, as Medicaid data quality is mixed and income data for commercial members may require linking in third-party data or use of unique data available for plans sold on the Health Insurance Marketplace.
  - Consensus was not reached on stratification structures beyond the Medicare line of business.

**Disability** and **primary language** were among the bottom three variables for both feasibility and importance.

## Recommended Stratification Structures

### DISABILITY

- **Current stratification structure:** Disabled / Non-Disabled (Medicare only)
- **Recommended stratification structure:** Disabled / Non-Disabled (Medicare only)
  - The TEP noted that within Medicare, it is only feasible to stratify disability based on whether an individual originally qualified for Medicare due to a disability. This is problematic, as individuals may develop a disability at later points.
  - The TEP was not confident in feasibility of consistent stratification by disability in Medicaid, citing varying eligibility requirements especially between states that did and did not expand Medicaid.
  - The TEP discussed use of a claims-based algorithm to more accurately identify disability, including in Medicaid and commercial populations, but conceded that it would entail substantial burden that reduces feasibility.
  - Consensus was not reached on stratification structures beyond the Medicare line of business.

### PRIMARY LANGUAGE

- **Current stratification structure:** Not included in standard PQA testing
- **Recommended stratification structure:** No standard structure identified
  - The TEP acknowledged the importance of English proficiency as a predictor of health outcomes. However, the TEP agreed that current feasibility is generally low due to variations in data structure and availability.
  - The TEP stressed the need for the healthcare system to find and consistently structure sources of reliable data for primary language to improve feasibility.
  - Consensus was not reached on a standard stratification structure due to the variability and limited availability of data.

Overall, TEP discussions for objective one were especially detailed as it was the first objective addressed by the TEP, and thus was the first set of discussions related to each of these stratification variables and structures. Additional details on TEP discussions for objective one are included in the Appendix.



## Next Steps

In alignment with TEP recommendations, **PQA has already piloted optional collection of race and ethnicity data** in its most recent testing plans. While data quality and completeness has been mixed, this represents an important step forward.

Many variables prioritized by the TEP are already included in PQA's testing plans. For these variables, PQA will explore updates to stratification structures in response to TEP recommendations (e.g., evaluation of additional age subgroups on a measure-by-measure basis). PQA will also explore the addition of prioritized variables that are not already included in testing plans; however, their lower feasibility and importance may pose challenges. The addition of lower-priority variables to testing plans must also balance the value of data with the added burden to testers.

## Objective 2 | Stratification of PQA Measures in Quality Programs

---

As a measure steward, PQA maintains strong relationships with quality program administrators like CMS and partners with them to implement PQA measure specifications. As part of this process, program administrators often make minor adjustments to measure specifications to allow a measure to better fit program logistics, data considerations, populations, and other needs. Program administrators also decide how measure rates are reported, both internally to measured entities and to the public.

The TEP was tasked with recommending stratification variables for PQA measures used in quality programs. These recommendations, which will be provided to program administrators, are advisory in nature and are provided as a **prioritized list of variables** for each program, with a particular focus on feasibility and immediate readiness. They are not prescriptive as to whether stratified rates are publicly reported or used internally by measured entities.

Before discussing individual programs, the TEP evaluated whether stratification should be implemented in quality programs even when initial measure testing does not suggest evidence of disparities. The TEP acknowledged that healthcare processes and policies can shift over time, meaning disparities that may not be present during testing could emerge later.<sup>8</sup> Consequently, the TEP agreed on recommending stratification in quality programs even when testing does not show disparities during development.

## Medicare Part D Star Ratings

### *Program Background*

CMS annually publishes Star Ratings for Medicare Part D Prescription Drug Plans (PDPs) and Medicare Advantage (Medicare Part C) Prescription Drug plans (MAPDs) to provide quality and performance information to Medicare

beneficiaries. These ratings, ranging from one to five stars, are released on the CMS Plan Finder website to assist consumers in choosing a high-quality health plan.<sup>9</sup> In addition to public reporting, Star Ratings performance has substantial financial impacts on plans, including MAPD quality bonus payments totaling over \$10 billion in 2023.<sup>10</sup>

Five PQA measures are included in the 2024 Star Ratings (2022 measurement year): *Proportion of Days Covered: Statins*, *Proportion of Days Covered: Diabetes All Class*, *Proportion of Days Covered: Renin-Angiotensin System Antagonists*, *Statin Use in Persons in Diabetes*, and *Comprehensive Medication Review (CMR) Completion Rate*. Two additional measures, *Polypharmacy: Use of Multiple Anticholinergic Medications in Older Adults* and *Concurrent Use of Opioids and Benzodiazepines (COB)* are scheduled to be added to the 2027 Star Ratings (2025 measurement year). Several additional PQA measures are also used in other Medicare Part D quality programs, including the Part D Display Page, which publicly reports measures but does not include financial incentives, and the Part D Patient Safety reports, which report measures confidentially to health plans.

Starting in the 2027 Star Ratings (2025 measurement year), the Health Equity Index (HEI) will be introduced to the Part C and D Star Ratings and will provide a bonus to plans for improved performance on patients with specific social risk factors (SRFs).<sup>7</sup> The SRFs included in the HEI include disability, dual-eligibility, and low-income subsidy (LIS). The HEI is separate from (and complementary to) the Categorical Adjustment Index, which also adjusts health plan Star Ratings based on a contract's percentage of patients with the same SRFs.

## TEP Discussions and Recommendations

**TABLE 3** | Prioritized List of Recommended Stratification Variables for PQA Measures in Medicare Part D Star Ratings

| VARIABLE         | MEDIAN | MEAN | MINIMUM | MAXIMUM | PERCENTAGE RANKING TOP 3 |
|------------------|--------|------|---------|---------|--------------------------|
| AGE              | 1      | 1.2  | 1       | 3       | 100%                     |
| SEX              | 2      | 2.4  | 1       | 3       | 100%                     |
| GEOGRAPHY        | 3      | 3.1  | 1       | 5       | 61%                      |
| RACE/ETHNICITY   | 4      | 3.5  | 1       | 5       | 39%                      |
| PRIMARY LANGUAGE | 5      | 4.8  | 4       | 5       | 0%                       |

*Note: Variables included in the HEI were not included in the prioritization exercise.*

Table 3 provides the prioritized list of recommended stratification variables for PQA measures used in the Medicare Part D Star Ratings. The TEP felt that most of their previous considerations on these variables in the context of PQA measure testing also applied in the context of quality programs. The TEP considered disability, dual-eligibility, and LIS status to be de-facto stratification priorities due to their inclusion in the HEI; consequently, these variables were not included in the prioritization exercise.

The TEP considered the quality of race and ethnicity data available in Medicare Part D, citing concerns about missingness and potentially inaccurate conclusions drawn by the algorithm Medicare Part D uses to impute certain race/ethnicities based on surname. However, the TEP agreed there is still value in the race and ethnicity data as it stands and recommended that CMS use what is currently available even if the data are not perfect. The TEP also supported aligning race and ethnicity stratification with CMS requirements for race and ethnicity data on enrollment data.

## Medicaid Adult Core Set

### Program Background

The Medicaid Adult Core Set tracks and publishes quality measures at the state level to enable comparisons, drive quality improvement, and evaluate the national quality of care for adult Medicaid beneficiaries. Annual ratings are publicly reported on Medicaid.gov.<sup>11</sup> PQA measures currently included in the Adult Core Set are *Concurrent Use of Opioids and Benzodiazepines (COB)* and *Use of Opioids at High Dosage in Persons Without Cancer (OHD)* (OHD will be retired from this program in 2026).

### TEP Discussions and Recommendations

**TABLE 4** | Prioritized List of Recommended Stratification Variables for PQA Measures in the Medicaid Adult Core Set

| VARIABLE                  | MEDIAN | MEAN | MINIMUM | MAXIMUM | PERCENTAGE RANKING TOP 3 |
|---------------------------|--------|------|---------|---------|--------------------------|
| AGE                       | 1      | 1.4  | 1       | 5       | 94%                      |
| SEX                       | 3      | 3.0  | 2       | 5       | 72%                      |
| GEOGRAPHY                 | 3      | 4.1  | 1       | 8       | 44%                      |
| DUAL ELIGIBILITY          | 4      | 4.4  | 2       | 9       | 33%                      |
| MEDICAID EXPANSION STATES | 5      | 5.1  | 1       | 9       | 28%                      |
| RACE/ETHNICITY            | 5.5    | 5.3  | 1       | 9       | 22%                      |
| INCOME                    | 7      | 6.7  | 3       | 9       | 6%                       |
| PRIMARY LANGUAGE          | 7      | 7.4  | 5       | 9       | 0%                       |
| DISABILITY                | 8      | 7.6  | 4       | 9       | 0%                       |

Table 4 provides the prioritized list of recommended stratification variables for PQA measures in the Medicaid Adult Core Set.

The TEP noted that geography is connected to health outcomes and is especially important for its connection to resource availability in this population, although some questioned what would be expected of Medicaid programs and managed care plans upon determining that certain geographic areas are experiencing disparities. The TEP also questioned whether county-level data would yield too many small numbers for the medications captured in PQA measures used in Medicaid; the TEP noted that county-level data is feasible for more prevalent chronic conditions (e.g., diabetes, asthma), although rarer conditions may present issues.

The TEP noted the Core Set currently provides guidance that states should stratify their rates by sex, ethnicity, race, and geography. The guidance does not currently include age, which the TEP felt was important to include.<sup>12</sup> The TEP noted that the stratification structure for geography as currently reported by the Core Set is urban, rural, missing, or other, but the Core Set does not define urban or rural. In addition, the TEP noted that stratification by reason for Medicaid eligibility could yield useful information for identifying disparities.

# Health Insurance Marketplace Quality Rating System

## Program Background

The Health Insurance Marketplace Quality Rating System (QRS) provides quality and performance information to individuals seeking coverage on the Health Insurance Marketplace to assist them in choosing qualified health plans.<sup>13</sup> Annual performance ratings are released on HealthCare.gov for consumers, and performance data are released on CMS.gov through a set of public use files. QRS ratings do not have financial incentives, although higher ratings may lead to increased enrollment. PQA measures included in the QRS are *Proportion of Days Covered: Diabetes All Class*, *Proportion of Days Covered: RASAs*, *Proportion of Days Covered: Statins*, *Annual Monitoring for Persons on Long-Term Opioid Therapy*, and *International Normalized Ratio Monitoring for Persons on Warfarin*.

## TEP Discussions and Recommendations

**TABLE 5** | Prioritized List of Recommended Stratification Variables Recommended for PQA Measures in the Marketplace Quality Rating System

| VARIABLE         | MEDIAN | MEAN | MINIMUM | MAXIMUM | PERCENTAGE RANKING TOP 3 |
|------------------|--------|------|---------|---------|--------------------------|
| AGE              | 1      | 1.3  | 1       | 3       | 100%                     |
| SEX              | 2      | 2.5  | 1       | 5       | 94%                      |
| GEOGRAPHY        | 3.5    | 3.3  | 1       | 6       | 50%                      |
| RACE/ETHNICITY   | 4.5    | 4.1  | 1       | 7       | 39%                      |
| INCOME           | 5      | 4.8  | 2       | 7       | 11%                      |
| PRIMARY LANGUAGE | 6      | 5.9  | 4       | 7       | 0%                       |
| DISABILITY       | 7      | 6.2  | 3       | 7       | 6%                       |

Table 5 provides the prioritized list of recommended stratification variables for PQA measures in the QRS.

The TEP noted that the Marketplace is growing in importance, as the most recent enrollment cycle saw the greatest growth in enrollment in program history, and discussed the use of metal level as an income-related variable.<sup>14</sup> A Marketplace plan's metal level is determined by its actuarial value, or the extent of essential health care costs projected to be paid by the plan versus the individual.<sup>15</sup>

TEP members supported the idea of using metal level as a proxy for income, noting that research has demonstrated a variety of disparities across metal levels. The TEP acknowledged that metal level is not a perfect variable, and exactly what it shows is not completely clear or consistent. However, the TEP agreed that it has value in identifying potential disparities in performance.

The TEP felt that stratifying by race and ethnicity, geography, and primary language was important in the Marketplace, although feasibility differed. The TEP also discussed digital access and digital literacy as potential influences on Marketplace measures; while digital access data are typically available in third party databases,

these data must be linked to individuals, which is a complex process. The TEP also raised the concern that many individuals who lose Medicaid coverage subsequently enroll in the Marketplace, and thus state Medicaid expansion status may be correlated with changes in Marketplace population characteristics.

Integrated Healthcare Association | Align. Measure. Perform. Program

Program Background

The Integrated Healthcare Association (IHA) Align. Measure. Perform. (AMP) program is a state-level program that monitors and evaluates the quality of health plan performance in the state of California.<sup>16</sup> Annual ratings are released by the California Office of the Patient Advocate to allow consumers to compare plans. Public reporting for this program is voluntary. PQA measures included in this program are *Proportion of Days Covered: Diabetes All Class*, *Proportion of Days Covered: RASAs*, *Proportion of Days Covered: Statins*, *Statin Use in Persons with Diabetes*, and *Concurrent Use of Opioids and Benzodiazepines*.

TEP Discussions and Recommendations

| TABLE 6   Prioritized List of Recommended Stratification Variables for the Integrated Healthcare Association AMP Program |        |      |         |         |                          |
|--|--------|------|---------|---------|--------------------------|
| VARIABLE   | MEDIAN | MEAN | MINIMUM | MAXIMUM | PERCENTAGE RANKING TOP 3 |
| AGE  | 1      | 1.3  | 1       | 3       | 100%                     |
| SEX  | 2      | 2.4  | 1       | 5       | 94%                      |
| GEOGRAPHY  | 4      | 3.4  | 1       | 6       | 44%                      |
| RACE/ETHNICITY   | 4      | 4.0  | 1       | 7       | 39%                      |
| INCOME   | 5.5    | 5.3  | 2       | 7       | 11%                      |
| PRIMARY LANGUAGE   | 6      | 5.6  | 3       | 7       | 6%                       |
| DISABILITY   | 6      | 5.9  | 3       | 7       | 6%                       |

Table 6 provides the prioritized list of recommended stratification variables for PQA measures in the Integrated Healthcare Association AMP program. TEP discussion specific to the IHA AMP program was limited, and the TEP supported aligning their recommendations with the other programs that were discussed.

Next Steps

PQA will provide the prioritized lists of recommended stratification variables to quality program administrators, who have been kept informed during development of this report. PQA anticipates that these recommendations will serve to be useful resources as quality programs continue to evaluate the role of stratification in their health equity strategies.

## Objective 3 |

### Affirming Existing PQA Adherence Risk Adjustment Model and Exploring Future Directions

---

#### Risk Adjustment Overview

Risk adjustment (or case-mix adjustment) is a statistical method used to account for patient-related factors (e.g., age, comorbidity, illness severity) that may impact health outcomes but may be outside the direct control of the measured entity (e.g., health plans, providers).<sup>4,6</sup> For example, a surgical mortality measure may be risk adjusted to account for the fact that certain surgeries carry greater risk of mortality. The goal of risk adjustment is to account for factors that are unrelated to the quality of care provided and improve the ability to make fair and accurate conclusions about the quality of care provided.

When applied in quality measurement, measure scores are adjusted based on identified factors and empirically determined weights to produce a single risk-adjusted performance score (in contrast to stratification, in which a performance score is calculated for each different subpopulation). From an equity perspective, risk adjustment can be important to avoid penalizing providers that care for disadvantaged patients and to evaluate true differences in quality of care rather than differences in case-mix of a population. In certain situations, a lack of risk adjustment can exacerbate or even institutionalize disparities when providers are penalized for factors outside of their control, which subsequently limits their access to the resources they need to serve their vulnerable populations and improve performance.

Risk adjustment for sociodemographic variables has received significant attention in recent years and has been subject to diverging and sometimes conflicting guidance. In its 2014 Technical Report for Risk Adjustment for Socioeconomic Status or Other Sociodemographic Factors, NQF recommended sociodemographic status (SDS) risk adjustment of performance measures if: (1) there is a conceptual relationship between SDS and the outcome of interest; and (2) there is empirical evidence that SDS affects the outcome of interest. SDS refers to a variety of socioeconomic (e.g., income, education, occupation) and demographic factors (e.g., age, sex, race, primary language) that may be associated with differential health outcomes.<sup>4,6</sup>

#### History of PQA Risk Adjustment of Adherence Measures

In response to NQF's recommendations, PQA convened the Risk Adjustment Advisory Panel (RAAP) in 2015 to determine which PQA measures may be appropriate for SDS risk adjustment, and to recommend a risk adjustment methodology for the measures.<sup>17</sup> The TEP was comprised of representatives from PQA's multistakeholder member organizations with expertise in risk adjustment methodology, medication adherence, health disparities research, and health plans currently participating in the CMS Part C and D Star Ratings program.

The RAAP identified three PQA-stewarded health plan measures that assess medication adherence by Proportion of Days Covered (PDC) as appropriate to consider for risk adjustment, in part because they are intermediate outcome measures; such measures are commonly (although not universally) considered more appropriate for risk adjustment compared to process measures.



PQA contracted with CMS to conduct a study assessing the impact of SDS variables on these three adherence measures. Through a systematic review of literature, discussion among SMEs, and voting, the RAAP selected SDS variables for risk adjustment and guided development of a valid risk adjustment model for the three identified PDC measures. PQA conducted analyses of a full model using beneficiary-level data from CMS, as well as additional community- and county- level data from other data partners (Table 7). However, given the challenges associated with obtaining SDS data, PQA also explored a reduced model that included only beneficiary-level factors that have data readily available to CMS and health plans.

Among beneficiary-level factors, the RAAP raised specific concerns regarding the race variable. In testing data, the quality and completeness of the race variable were lacking, with concerns noted about substantial missing data and issues with the integrity of the data which were available. Additionally, at the time, leading bodies like NQF were concerned about the use of race in risk adjustment models as a proxy for socioeconomic status, and the idea that including race creates a differential quality standard by race. As a result, PQA explored a reduced risk adjustment model both with and without race (i.e., analysis was limited to age, gender\*, low-income subsidy (LIS)/dual eligibility status, and disability status), and a comparison of the models showed similar risk-adjusted rates for the Part D contracts.

Overall, the results of the full and reduced models were similar for all three adherence measures with respect to both direction and strength of impact. Therefore, PQA recommended the reduced risk adjustment model.

*\*The term “gender” is used above and in the table and section below because it has been historically used in the context of this risk adjustment model, including in current documentation. Within this model and for the operational purposes of this report, “gender” is herein defined identically to “sex”.*

TABLE 7

Variables Included in SDS Analysis for Three PQA PDC (Adherence) Measures

| VARIABLE LEVEL                  | VARIABLE  | FULL MODEL | REDUCED MODEL |
|---------------------------------|---|------------|---------------|
| BENEFICIARY                     | Age   | •          | •             |
|                                 | Gender  | •          | •             |
|                                 | Low-Income Subsidy (LIS) status or Dual eligibility status        | •          | •             |
|                                 | Disability as original reason for Medicare entitlement            | •          | •             |
|                                 | Race  | •          |               |
| COMMUNITY<br>(9-DIGIT ZIP CODE) | Median income   | •          |               |
|                                 | Percent of households where residents are married                 | •          |               |
|                                 | Percent of households where residents completed college           | •          |               |
|                                 | Percent of households where residents own their home              | •          |               |
| COUNTY                          | Federally designated primary care professional shortage area      | •          |               |
|                                 | Federally designated mental healthcare professional shortage area | •          |               |

The PQA SDS risk adjustment model for the three adherence measures adjusts for the following patient factors: age, gender, LIS/dual eligibility status, and disability status. The SDS risk-adjusted PQA adherence measures are currently included in the 2024 Patient Safety Reports. CMS issued a final rule, published on April 12, 2023, finalizing implementation of PQA’s recommended methodology for SDS risk adjustment for the three adherence measures.<sup>7</sup> The final rule stated that the risk-adjusted measures will be included on the Part D Display Page for 2026 and 2027 (reporting on data from the 2024 and 2025 measurement years, respectively); they will move into the Part D Star Ratings beginning with the 2028 Star Ratings (reporting on data from the 2026 measurement year) while the unadjusted measures will be removed from the Star Ratings in the same year.

## TEP Discussions and Recommendations

The TEP was tasked with reviewing and affirming the risk adjustment model currently recommended for use with the adherence health plan measures used in the Medicare Part D Star Ratings. The TEP reviewed and discussed the history of PQA's adherence risk adjustment model, and briefly discussed the current state of the risk model while acknowledging the complexity of the implementation process and substantial time required to implement any future changes. Overall, they affirmed the importance and appropriateness of the current model while noting that the current risk adjustment model's implementation represents substantial progress and should be recognized as a success.

The TEP provided input on additional variables that could potentially be explored in future iterations of the risk adjustment model, especially in response to developments in data quality and availability or new information on health equity. The variables most frequently mentioned were **race** and **ethnicity**, as the TEP expected quality and completeness of these data to improve over time.

Other proposed variables for future exploration were **income or other indicators of economic status** and **geography/geographic access**. TEP members cited evidence that income is associated with higher healthcare resource utilization and inability to pay is often a primary reason for medication nonadherence. Geography or geographic access, classified as urban, rural, or suburban, was also mentioned as a meaningful variable to explore given that differences in adherence may be associated with geographic factors.

The TEP suggested exploring a variable that captures **household size or availability of social support**, noting that loneliness and social isolation are shown to result in poorer health outcomes and onset of chronic disease. Household data from the U.S. Census Bureau, including the American Community Survey,<sup>18</sup> includes information on "single-person households" which could potentially be leveraged as an indicator for household size.

Other proposed variables for exploration included **English language proficiency**, **presence of SDOH-related Z-codes**,<sup>19</sup> and a **community-level index** such as Area Deprivation Index (ADI),<sup>20</sup> Social Vulnerability Index (SVI),<sup>21</sup> or Minority Health Social Vulnerability Index.<sup>22</sup> The group qualified that the data needed for these variables may be difficult to capture due to low usage (e.g., for documentation of Z-codes) or lack of required reporting.

Overall, TEP discussion affirmed the value of the current risk adjustment model while noting that the substantial complexity of implementing changes in programs like Medicare Part D makes updating the risk adjustment model challenging. Beyond implementation challenges, updates to the risk adjustment model are also cost- and resource-intensive. However, the TEP ultimately offered valuable insight for potential future iterations of the PQA adherence risk adjustment model.

## Next Steps

Given the TEP's approval of the current PQA adherence risk adjustment model and its imminent implementation in the Medicare Part D Star Ratings, no action is currently needed. The TEP's insight on potential additional variables may be considered if future iterations of the risk adjustment model are developed.

# Objective 4 |

## Feasibility of Risk Adjusting the PDC-CMP Health Plan Measure

---

### Background

The PDC-CMP measure was developed in response to stakeholder support for a summary indicator of three PQA adherence measures (Statins, Diabetes All Class, RAS Antagonists). Additionally, the creation of the PDC-CMP measure was anticipated to improve continued usability of PDC measures in programs such as the Medicare Part D Star Ratings over time. While performance on PDC measures has improved continuously in the Part D program, PQA has not seen evidence that performance has “topped out”, or reached such a high level that there is no sufficient room for improvement or differentiation between entities. In the future, when performance does begin to near a top out, PQA envisions the potential use of the PQA composite as a valuable roll-up of measures to ensure continued focus while also reducing the overall number of measures in quality programs.

PQA typically evaluates the need for risk adjustment during the initial measure development process, but pursues development of the risk adjustment model after endorsement due to the substantial burden and time associated with building and testing the model. Due to the SDS risk adjustment recommended for the individual PDC measures, PQA also identified risk adjustment as potentially appropriate for the PDC-CMP measure during development.

Risk adjustment of a composite measure is technically complex. The PDC-CMP measure is structured at the person-measure level, (e.g., a person is a separate numerator and denominator case for the statin, diabetes, and RASA components), which means an individual may be counted multiple times in the measure by falling into multiple components. This complicates common risk adjustment methodologies, which evaluate typically patient-level factors and can therefore be impacted by patients counting more than once.

### Subject Matter Expert and TEP Discussions and Recommendations

PQA engaged a statistical SME to gain more insight into the development of a risk adjustment model for the PDC-CMP measure. The SME advised that in order to account for multiple observations per patient, a three-level model may be conceptually appropriate, nesting observations at the contract, person, and measure-component level. The SME also noted that there is existing evidence describing a small difference between this three-level approach and the approach used within the risk adjustment model for the individual PDC measures (ignoring the issue of multiple observations per patient). However, regardless of the approach taken, development of a risk adjustment model for the PDC-CMP measure would require significant time and resources.

This information was presented to the TEP. With an understanding that developing a model may be feasible in the future, PQA requested recommendations on which variables to consider for inclusion in a future model. In particular, beyond the inclusion of variables discussed in the previous section, PQA asked whether there were any conceptual considerations PQA should be aware of for the PDC-CMP measure in comparison to individual adherence measures.

The TEP was aligned with using the variables currently used in existing adherence risk adjustment model and additional variables recommended in the previous section as a starting point for the PDC-CMP measure, including **age, sex, disability, dual eligibility** and **LIS**, and **race and ethnicity** among others.

The TEP also raised questions about differential weighting of certain components; for example, adherence to diabetes medications could more significantly impact disadvantaged populations. PQA noted that during development of the PDC-CMP measure, consideration was given to weighting the components differently, although no weighting was ultimately used due to lack of sufficient rationale to weight a given component higher than another based on empirical evidence or differential performance between components. The TEP noted that programs like the Part D Star Ratings compare plans to one another, so all risk adjustment is relative.

## Next Steps

Given the resources required, future development of a risk adjustment model for the PDC-CMP measure will likely be contingent on use of the measure in a major quality program. PQA will use the Health Equity TEP's recommendations as the starting point if and when risk adjustment is explored for the PDC-CMP measure.

# Conclusions and Considerations for Future Work

---

The TEP provided valuable recommendations that serve as a roadmap to more equitably enhance medication use quality measurement. PQA encourages users of PQA measures, such as CMS and health plans, to leverage these recommendations to explore where disparities may exist in their own performance. These recommendations can also benefit researchers seeking to integrate health equity into their investigations, educators who can include it in their measurement-related materials, and a host of other stakeholders.

This work aligns well with NQF's recommendation to build a conceptual model illustrating the connections between social risk factors, clinical risk factors, the clinical process of interest, and associated outcomes.<sup>2</sup> The TEP's recommendations, which consider the connections between social risk factors and processes and outcomes of interest, provide a strong rationale for stratification structures.

TEP discussions frequently underscored the importance of continued efforts to standardize definitions and structures of sociodemographic data. Despite numerous efforts towards standardization to date, both from CMS<sup>23</sup> and from industry stakeholders such as National Council for Prescription Drug Programs,<sup>24</sup> many types of sociodemographic data are still variably defined in practice across stakeholders. This variation limits consistent stratification and inhibits the validity of comparisons between entities. Recognizing this issue, the CMS Framework for Health Equity's first objective is to Expand the Collection, Reporting, and Analysis of Standardized Data across federal programs and agencies.<sup>25</sup> The findings of this TEP underscore the importance of this objective.

In keeping with the theme of standardization, measure developers and stewards should strive to harmonize their approaches to measure stratification with these and other recommendations. Consistent stratification across measures from different sources will allow better comparisons and conclusions about the quality of care offered to vulnerable subgroups.

Finally, PQA recognizes and emphasizes that identification of disparities is not enough to achieve health equity. Quality measurement is a critical tool to highlight existing disparities and track progress, but it must be accompanied by investments and interventions to address gaps and improve care provided to vulnerable subgroups. PQA is supportive of efforts that align incentives with improvement on health equity-related measurement, such as the HEI, and looks forward to future efforts with its members and the broader healthcare system to advance health equity.

# About PQA

---



PQA, the Pharmacy Quality Alliance, is a national quality organization dedicated to improving medication safety, adherence and appropriate use. A measure developer, researcher, educator and convener, PQA's quality initiatives support better medication use and high-quality care. PQA was established in 2006 as a public-private partnership with the Centers for Medicare & Medicaid Services. PQA was created because prescription drug programs were a major area of health care where there was no organization or national program focused on quality improvement. Today, PQA is an independent, non-profit organization with more than 200 diverse members across health care.

## **Report Authors:**

Ben Shirley, CPHQ, *PQA Senior Director, Performance Measurement*

Melissa Castora-Binkley, PhD, *PQA Senior Director, Research*

Carly Vogel, PhD, *PQA Research Associate*

Razanne Oueini, PharmD, MSc, CPHQ, *PQA Senior Manager, Performance Measurement*

Lisa Hines, PharmD, CPHQ, *PQA Chief Quality and Innovation Officer*

***For questions or inquiries about this report or the PQA Health Equity Technical Expert Panel, contact PQA at [MeasureUse@PQAalliance.org](mailto:MeasureUse@PQAalliance.org).***



# References

---

1. U.S. Department of Health and Human Services. *2023 Equity Action Plan Summary: U.S. Department of Health and Human Services*. 2023. Accessed July 16, 2024. <https://www.whitehouse.gov/wp-content/uploads/2024/02/HHS-2023-EAP-Summary.pdf>
2. National Quality Forum. *Developing and Testing Risk Adjustment Models for Social and Functional Status-Related Risk Within Healthcare Performance Measurement: Final Technical Guidance – Phase 2*. 2022. Accessed January 23, 2023. [https://www.qualityforum.org/Projects/n-r/Risk\\_Adjustment\\_Guidance/Technical\\_Guidance\\_Final\\_Report\\_-\\_Phase\\_2.aspx](https://www.qualityforum.org/Projects/n-r/Risk_Adjustment_Guidance/Technical_Guidance_Final_Report_-_Phase_2.aspx)
3. U.S. Department of Health and Human Services. *Report to Congress: Social Risk Factors and Performance Under Medicare’s Value-Based Purchasing Programs*. 2016. Accessed February 14, 2023. <https://aspe.hhs.gov/reports/report-congress-social-risk-factors-performance-under-medicare-value-based-purchasing-programs>
4. U.S. Department of Health and Human Services. *Executive Summary - Report to Congress: Social Risk Factors and Performance in Medicare’s Value-Based Purchasing Program*. 2020. Accessed January 23, 2023. <https://aspe.hhs.gov/reports/second-report-congress-social-risk-medicare-value-based-purchasing-programs>
5. National Quality Forum. *Risk Adjustment for Socioeconomic Status or Other Sociodemographic Factors: Technical Report*. 2014. Accessed February 28, 2023. [https://www.qualityforum.org/Publications/2014/08/Risk\\_Adjustment\\_for\\_Socioeconomic\\_Status\\_or\\_Other\\_Sociodemographic\\_Factors.aspx](https://www.qualityforum.org/Publications/2014/08/Risk_Adjustment_for_Socioeconomic_Status_or_Other_Sociodemographic_Factors.aspx)
6. National Quality Forum. *Developing and Testing Risk Adjustment Models for Social and Functional Status-Related Risk Within Healthcare Performance Measurement: Final Technical Guidance – Version 4*. 2021. Accessed February 28, 2023. [https://www.qualityforum.org/Publications/2021/08/Developing\\_and\\_Testing\\_Risk\\_Adjustment\\_Models\\_for\\_Social\\_and\\_Functional\\_Status-Related\\_Risk\\_Within\\_Healthcare\\_Performance\\_Measurement\\_-\\_Final\\_Technical\\_Guidance.aspx](https://www.qualityforum.org/Publications/2021/08/Developing_and_Testing_Risk_Adjustment_Models_for_Social_and_Functional_Status-Related_Risk_Within_Healthcare_Performance_Measurement_-_Final_Technical_Guidance.aspx)
7. Centers for Medicare & Medicaid Services. Medicare Program; Contract Year 2024 Policy and Technical Changes to the Medicare Advantage Program, Medicare Prescription Drug Benefit Program, Medicare Cost Plan Program, and Programs of All-Inclusive Care for the Elderly. 2023. p. 22120-22345.
8. 117th United States Congress. H.R.5376 - Inflation Reduction Act of 2022. Updated Aug 16, 2022. <https://www.congress.gov/bill/117th-congress/house-bill/5376/text>.
9. Centers for Medicare & Medicaid Services. 2024 Medicare Advantage and Part D Star Ratings. U.S. Department of Health and Human Services. Accessed August 23, 2024. <https://www.cms.gov/newsroom/fact-sheets/2024-medicare-advantage-and-part-d-star-ratings>
10. Biniek JF, Damico A, Neuman T. Spending on Medicare Advantage quality bonus payments will reach at least \$12.8 billion in 2023. KFF. Accessed August 2, 2024. <https://www.kff.org/medicare/issue-brief/spending-on-medicare-advantage-quality-bonus-payments-will-reach-at-least-12-8-billion-in-2023/>
11. Centers for Medicare & Medicaid Services. Adult health care quality measures. Accessed August 20, 2024. <https://www.medicare.gov/medicaid/quality-of-care/performance-measurement/adult-and-child-health-care-quality-measures/adult-health-care-quality-measures/index.html>

12. Centers for Medicare & Medicaid Services. *Technical Assistance Resource: Reporting Stratified Results in the Quality Measure Reporting System for the 2024 Child, Adult, and Health Home Core Sets*. 2024. Accessed October 8, 2024. <https://www.medicare.gov/medicaid/quality-of-care/downloads/QMR-stratification-resource.pdf>
13. Centers for Medicare & Medicaid Services. About the Quality Rating System (QRS). U.S. Department of Health and Human Services. Updated June 28, 2024. Accessed August 23, 2024. <https://www.cms.gov/Medicare/Quality-Initiatives-Patient-Assessment-Instruments/QualityInitiativesGenInfo/ACA-MQI/Quality-Rating-System/About-the-QRS>
14. Kaiser Family Foundation. Marketplace Enrollment, 2014-2024. Accessed October 9, 2024. <https://www.kff.org/affordable-care-act/state-indicator/marketplace-enrollment/?currentTimeframe=0&sortModel=%7B%22colId%22:%22Location%22,%22sort%22:%22asc%22%7D>
15. Levitt L, Claxton G. *What the Actuarial Values in the Affordable Care Act Mean*. 2011. Accessed August 20, 2024. <https://www.kff.org/wp-content/uploads/2013/01/8177.pdf>
16. Integrated Healthcare Association. Align. Measure. Perform. Program. Accessed August 23, 2024. <https://iha.org/performance-measurement/amp-program/>
17. Pharmacy Quality Alliance. National Quality Forum Endorses PQA's Risk Adjusted Adherence Measures. Pharmacy Quality Alliance. Accessed February 28, 2023. <https://www.pqaalliance.org/nqf-endorses-pqa-risk-adjusted-measures>
18. United States Census Bureau. American Community Survey (ACS). Updated June 10, 2024. Accessed August 23, 2024. <https://www.census.gov/programs-surveys/acs>
19. Centers for Medicare & Medicaid Services. Improving the collection of Social Determinants of Health (SDOH) data with ICD-10-CM Z codes. Accessed August 23, 2024. <https://www.cms.gov/files/document/cms-2023-omh-z-code-resource.pdf>
20. Center for Health Disparities Research University of Wisconsin School of Medicine and Public Health. Neighborhood Atlas. Applied Population Lab University of Wisconsin-Madison. Accessed August 23, 2024. <https://www.neighborhoodatlas.medicine.wisc.edu/>
21. Centers for Disease Control and Prevention and Agency for Toxic Substances and Disease Registry. CDC/ATSDR Social Vulnerability Index (CDC/ATSDR SVI): Overview. U.S. Department of Health and Human Services. Updated June 14, 2024. Accessed August 19, 2024. <https://www.atsdr.cdc.gov/place-health/php/svi/index.html>
22. U.S. Department of Health and Human Services Office of Minority Health. Minority Health SVI. Updated June 2023. Accessed August 20, 2024. <https://minorityhealth.hhs.gov/minority-health-svi>
23. Henry J, Mekler S. ONC health IT framework for advancing SDOH data use and interoperability. Assistant Secretary for Technology Policy. Accessed October 22, 2024. [https://www.healthit.gov/buzz-blog/interoperability/onc-health-it-framework-for-advancing-sdoh-data-use-and-interoperability?utm\\_source=feedburner&utm\\_medium=feed&utm\\_campaign=Feed%3A+healthitbuzz-blog+%28Health+IT+Buzz+Blog%29](https://www.healthit.gov/buzz-blog/interoperability/onc-health-it-framework-for-advancing-sdoh-data-use-and-interoperability?utm_source=feedburner&utm_medium=feed&utm_campaign=Feed%3A+healthitbuzz-blog+%28Health+IT+Buzz+Blog%29)
24. National Council for Prescription Drug Programs. Strategic Initiatives. Accessed October 30, 2024. <https://www.ncpdp.org/Strategic-Initiatives.aspx>
25. Centers for Medicare & Medicaid Services. CMS Framework for Health Equity. U.S. Department of Health and Human Services. Updated September 10, 2024. Accessed October 17, 2024. <https://www.cms.gov/priorities/health-equity/minority-health/equity-programs/framework>

26. U.S. Department of Health and Human Services Office of Minority Health. Explanation of data standards for race, ethnicity, sex, primary language, and disability. Accessed August 23, 2024. <https://minorityhealth.hhs.gov/explanation-data-standards-race-ethnicity-sex-primary-language-and-disability>
27. Office of Management and Budget. Revisions to OMB's Statistical Policy Directive No. 15: Standards for Maintaining, Collecting, and Presenting Federal Data on Race and Ethnicity. 2024. p. 22182-22196.
28. Embers Project. Community Needs Index. Edge Collaborative. Accessed August 19, 2024. <https://www.embersproject.org/cni>
29. Health Resources & Services Administration. What is shortage designation? Accessed August 19, 2024. <https://bhwh.hrsa.gov/workforce-shortage-areas/shortage-designation>
30. Quest Analytics. Quest Analytics to measure adequate access to care for Medicare Advantage, Medicare-Medicaid and Pharmacy Plans. PR Newswire. Accessed October 7, 2024, 2024. <https://www.prnewswire.com/news-releases/quest-analytics-to-measure-adequate-access-to-care-for-medicare-advantage-medicare-medicaid-and-pharmacy-plans-301130692.html>
31. Centers for Medicare & Medicaid Services. Limited income and resources. U.S. Department of Health and Human Services. Accessed August 15, 2023. <https://www.cms.gov/training-education/partner-outreach-resources/low-income-subsidy-lis>
32. Centers for Medicare & Medicaid Services. DQ Atlas: Explore by topic, income. Department of Health and Human Services. Accessed July 31, 2023. <https://www.medicaid.gov/dq-atlas/landing/topics/single/map?topic=g3m17&tafVersionId=38>
33. Centers for Medicare & Medicaid Services. Coordination of benefits. U.S. Department of Health and Human Services. Accessed August 19, 2024. <https://www.healthcare.gov/glossary/coordination-of-benefits/#:~:text=A%20way%20to%20figure%20out,paying%20the%20same%20medical%20claim>
34. Research Data Assistance Center. Original Reason for Entitlement Code. Accessed August 23, 2024. <https://resdac.org/cms-data/variables/original-reason-entitlement-code>
35. Social Security Administration. Social Security Administration (SSA) Quarterly data for spoken language preferences: Retirement and Survivor Insurance (RSI) initial claims. Updated July 9, 2024. Accessed August 19, 2024. <https://www.ssa.gov/open/data/LEP-Quarterly-Spoken-Language-RSI-Claimants.html>
36. Centers for Medicare & Medicaid Services. *Contract Year 2021 Translated Model Materials Requirements and Language Data Analysis Methodology*. 2020. Accessed September 17, 2024. <https://www.hhs.gov/guidance/document/contract-year-2021-translated-model-materials-requirements-and-language-data-analysis>

# Appendix |

## TEP Discussion Details on Stratification Variables for PQA Measure Testing

---

### AGE

Stratifying measure rates by age is currently included in PQA testing plans with a standard stratification structure of 18-50, 51-64, 65-84, and 85+. The TEP agreed on the importance of stratifying by age as an equity-related variable, expressing that age can play different roles in an individual's health at different times in life. PQA noted that while the National Quality Forum has often considered age to be a clinical rather than a social variable, it has a role for both.

Regarding the stratification structure, TEP members discussed various approaches. TEP members noted that the current bracket of 18-50 years is a larger range than others, and suggestions for reducing this range included separate brackets for 18-34 and 35-50, as well as by brackets by generation (e.g., millennials versus baby boomers). TEP members also discussed the need to consider reducing the 65-84 range, noting that many subgroups within this range have distinct characteristics. For example, some segments of the 65-84 range may include more individuals who are still working versus retired, and younger beneficiaries who are currently aging into Medicare may be more technologically savvy than previously aged-in Medicare beneficiaries. Overall, TEP members agreed there is benefit to further breaking out the 18-50 and 65-84 age brackets.

Survey results demonstrated TEP consensus on using the following age categories: 18-34, 35-49, 50-64, 65-74, 75-84, and 85+. The TEP considered this stratification structure a starting point and encouraged tailoring it by measure as appropriate based on clinical factors or measure focus.

### SEX

Stratification by sex is currently included in PQA testing plans with a standard stratification structure of male, female. These data, in this format, are nearly universally available within insurance enrollment data.

The TEP agreed that sex has clinical and social components and is a distinct concept from sexual orientation and gender identity (SOGI). TEP members noted there are numerous ways to ask questions related to gender and sexual orientation, and these questions are frequently posed to patients in different ways by various entities. The TEP expressed that in an ideal world, PQA would stratify by SOGI variables; however, data on SOGI are rarely available, and a more consistent approach to collecting SOGI data is needed before stratification can occur in a uniform manner. The TEP conceded that stratifying by sex in a binary fashion is the most feasible option based on how data are currently collected.

Survey results demonstrated TEP consensus on use of a male, female stratification structure for sex in PQA measure testing, although comments and discussion demonstrated preference for continued evaluation of the feasibility of additional options and SOGI variables.

## RACE AND ETHNICITY

At the time of TEP discussion, PQA testing plans did not stratify by race and ethnicity. Race and ethnicity data are routinely available in Medicare data, although the quality of the data is mixed. However, race and ethnicity data are frequently missing in Medicaid data and are highly inconsistent in commercial data.

The TEP discussion largely focused on use of the minimum categories for race and ethnicity provided by the 1997 OMB Standards for Maintaining, Collecting, and Presenting Federal Data on Race and Ethnicity, which are listed below.<sup>26</sup> TEP members highlighted the importance of race and ethnicity, citing their strong associations with disparate disease progression, care processes, and outcomes. TEP members universally agreed that race and ethnicity variables are critical and including them in stratification may help encourage more consistent data collection by health plans and data aggregators. In discussing various ways to categorize race and ethnicity, TEP members acknowledged the need to balance the desire to be specific while keeping the categories broad enough to conduct meaningful analyses and avoid tester burden. TEP members noted that increasingly specific strata of race and ethnicity may raise sample size concerns as fewer and fewer patients fall into each specific strata.

TEP members also discussed sources of race and ethnicity data. TEP members noted that ideally, PQA measure testers would identify whether their data represents race and ethnicity collected via an administrative method or data that was self-reported; self-reported data is typically considered the gold standard for race and ethnicity. TEP members, especially patient representatives, also noted that inclusion of some standard option representing multiple races is ideal. They noted that many people identify as biracial and without an explicit option, they may be unsure of what to report. However, Medicare often relies on a combination of social security data and statistical algorithms as opposed to member-reported race, and use of self-reported data is rare in Medicaid and unknown in commercial data.

Survey results demonstrated TEP consensus on using the OMB minimum categories to structure race and ethnicity data in PQA measure testing.

OMB Standards minimum categories:<sup>26</sup>

- Ethnicity Categories
  - Hispanic or Latino
  - Not Hispanic or Latino
- Race Categories (minimum)
  - American Indian or Alaska Native
  - Asian
  - Black or African American
  - Native Hawaiian or Other Pacific Islander
  - White

In March 2024, after TEP discussions on race and ethnicity had concluded, OMB released revisions to Standards for Maintaining, Collecting, and Presenting Federal Data on Race and Ethnicity for the first time since 1997.<sup>27</sup> Importantly for the TEP recommendations, these revisions added Middle Eastern or North African as a new minimum race category. They also combined race and ethnicity data collection into a single question and encouraged respondents to select as many options as apply to how they identify. PQA will align with this new guidance in the future when possible.

## GEOGRAPHY

PQA testing plans do not currently stratify by geography. TEP members noted that geography is an important characteristic that encompasses a broad variety of underlying aspects across access, environmental health, and others. TEP discussion centered on the challenge of how best to consistently define geographic stratification in PQA testing. A variety of potential approaches were raised, including use of various indices such as Community Needs Index score,<sup>28</sup> Health Resources & Services Administration Health Professional Shortage Areas (HPSAs),<sup>29</sup> and others. The approach of urban versus rural was supported for simplicity, although some preference for including a suburban classification was also expressed.

TEP members noted many different options for defining areas as urban or rural. Mapping beneficiaries' ZIP codes to an urban/rural classification was discussed most, although there are different taxonomies available for this mapping. The TEP discussed that health plans currently assess beneficiary access for Medicare Advantage network adequacy through a tool provided by a CMS-contracted vendor,<sup>30</sup> which already includes urban/rural as well as suburban designations.

PQA noted that mapping ZIP codes to urban or rural designations would likely be burdensome to measure testers. The mapping would need to be completed by the testing partners based on PQA instruction since results provided by testing partners are summary-level and not beneficiary-level.

The TEP did not come to a final consensus regarding the recommended stratification structure for geography, especially given the multiple different methodologies available for defining urban, rural, and suburban. While their guidance provides a valuable starting-point, it underscores the current challenges and need for broader healthcare system consensus on how these terms should be defined and operationalized.

## DUAL-ELIGIBLE STATUS

Stratification by dual-eligible status (i.e., eligibility for both Medicare and Medicaid) is currently included in PQA testing plans, and these data are typically available from health plan testing partners. PQA's stratification structure for this variable reflects how source data are usually structured: dually eligible or non-dually eligible. TEP members agreed on the importance of including dual eligibility for stratification.

One member noted that this variable is especially important given the upcoming changes to Medicare Star Ratings reflecting the Health Equity Index (HEI),<sup>7</sup> which was discussed in greater depth in the next section. TEP members also noted that including income along with dual-eligible status would enhance the value of the data and related learnings.

Given the straightforward nature of the dual-eligible stratification structure, TEP consensus was considered unnecessary and was not pursued for this variable.

## INCOME

Stratification by an income indicator is currently included in PQA testing plans only for Medicare using low-income subsidy (LIS) status. The LIS is a subsidy paid by the federal government to a Medicare Part D drug plan for beneficiaries who need extra help with their prescription drug costs due to limited income and resources.<sup>31</sup> The current stratification structure is LIS, non-LIS.



LIS is not included in Medicaid data. Medicaid eligibility files instead contain data on income level with public files classifying income in eight divisions. However, a Medicaid Data Quality Atlas analysis of 2022 data found that 22 states were missing income values for more than 10% of beneficiaries, and of these 22 states, 9 were missing income values for greater than 50% of beneficiaries.<sup>32</sup> This raises a considerable challenge for assessing income status in this population.

The TEP discussion focused on income data within commercial plans and the feasibility of requiring different types of income indicators for different lines of business. For individuals enrolled in the Health Insurance Marketplace, TEP members discussed how inferences could be made about income based on the health plan category, or “metal level” (e.g., gold, silver, bronze) of the plan the member selected. Metal levels are based on the plan’s share and the member’s share of costs for covered services, while cost-sharing and premium subsidies that scale with income may also provide insights. However, the TEP conceded that these variables are imperfect proxies for income.

The TEP noted that data from third-party vendors could be used to link at the member level for commercial plan, although the data is likely to be imperfect. Use of a coordination of benefits indicator, which is used to assist in determining payments for individuals enrolled in multiple insurance plans,<sup>33</sup> was also raised as a potential marker for income.

The TEP agreed that income stratification may aid in making fair comparisons between across plans. Notably, a patient representative commented that income was the number one external influence on health outcomes in their opinion. However, the TEP did not come to a final consensus regarding the recommended stratification structure of the income variable beyond continued use of LIS status in Medicare.

## DISABILITY

Stratification by disability is currently included in PQA testing plans in Medicare only, using the stratification structure of disabled or non-disabled. These data are based on a beneficiary’s original reason for entitlement code (OREC),<sup>34</sup> which is readily available within Medicare data. However, these data do not capture individuals who become disabled after Medicare enrollment.

The TEP called out that disability is one of the few factors used to calculate the HEI, which will impact the Medicare Part C & D Star Ratings and reward factors. For that reason, they believed stratification by disability needs to remain for PQA testing in Medicare data, at a minimum. While the TEP suggested that the OREC codes were an imperfect indicator of disability, they noted that it has the advantage of being readily available. The TEP discussed the possibility of using a validated claims-based algorithm to capture a more inclusive view of disability, but also acknowledged this approach would be a considerable burden both to PQA to develop or adapt an algorithm, and to testers to apply it.

The TEP expressed concerns with the quality of disability data for Medicaid, noting that depending on each state’s eligibility criteria, individuals may become eligible for Medicaid for a reason other than disability even if they are currently disabled. Some TEP members expressed concerns about the evolving meaning of disabilities and how organizations may variably define and collect that information in the commercial line of business. The TEP noted that disability status can shift over time, and those changes may not be easily captured in the data.

The TEP did not come to consensus regarding the recommended stratification structure of the disability variable beyond use of continued use of OREC-derived disabled, non-disabled values in Medicare.

## PRIMARY LANGUAGE

Stratification by primary language is not currently included in PQA testing plans.

While some population-level data on primary language preferences in Medicare are available through the Social Security Unified Measurement System,<sup>35</sup> beneficiary-level primary language data is not currently available in Medicare data. While primary language is available in Medicaid, there are varying structures and missing data is frequent.

The TEP discussion focused on the feasibility of stratifying by primary language given the heterogeneity in source data across Medicaid programs and uncertain access among Medicare and commercial plans. The TEP noted that CMS provides various language-related data to plans, including marking geographic areas where with significant populations of non-English language speakers where certain vital materials must be translated upon beneficiary request.<sup>36</sup> The TEP also discussed how consistent use of interpreters during healthcare encounters has an important impact on the role of primary language in quality of care.

Overall, many TEP members agreed that language data has importance while a few members disagreed, noting that primary language could be a proxy or overlap with other variables that are more readily available. TEP members universally agreed that language data are difficult to obtain, and may need to be deprioritized compared to other discussed variables. The TEP did not come to a final consensus regarding the recommended stratification structure of the primary language variable.

## ADDITIONAL VARIABLES

The prioritization survey also included an opportunity for TEP members to submit other variables for future consideration. Marital status and/or household size were raised as potential proxies for living alone, caregiver support, loneliness, and more. Education level (e.g., high school, college) was also raised; the TEP noted data availability and feasibility is currently low, but consideration could be given to area-level health equity indices like the Social Vulnerability Index (SVI).<sup>21</sup> For each of these variables, additional research is needed for further consideration.