NATIONAL QUALITY FORUM
Driving measurable health improvements together

Current Practices of Testing Social and Functional Status-Related Risk Factors within Risk Adjustment Models of Performance Measurement

ENVIRONMENTAL SCAN DRAFT REPORT – VERSION 2

January 22, 2020

This report is funded by the Centers for Medicare & Medicaid Services under contract HHSM-500-2017-00060I HHSM-500-0000.

Table of Contents

Executive Summary

Since healthcare outcomes are a function of patient attributes, as well as the care received, and since patients are not randomly assigned to providers for healthcare services, risk adjustment is essential to ensuring valid comparisons between providers when examining outcome performance in real-world settings. There is a large body of evidence that various social and functional status-related factors influence outcomes, and thus influence results on outcome performance measures. However, measure developers have long expressed a need for technical guidance on developing and testing social and/or functional risk adjustment models for measure endorsement.

Building on several years of work with developing guidance for risk adjustment model development, the National Quality Forum (NQF) convened a technical expert panel (TEP) to work towards consensus decisions that will yield technical guidance for measure developers that includes emerging good and best practices on when and how to adjust for functional and social risk factors in measure development. To inform this guidance, NQF conducted a TEP-informed environmental scan of data sources used for risk adjustment, functional or social risk factors available for testing, and approaches to conceptual and statistical methods for risk adjustment.

Within this report various data sources and testing approaches were identified for social and functional factor risk adjustment. Themes were identified in the types of data and methods used within the literature and within the illustrative set of measures submitted for NQF endorsement. Administrative claims, registry data, and electronic health records remain the primary data sources used for quality measurement development. Social factors included for risk adjustment analyses are largely at the patient- and community-levels, with the latter being sourced from various socioeconomic and sociodemographic indices. Functional risk factors identified were all at the patient-level, as these data tend to be captured directly from the patient through survey instruments and/or assessments. Compared to social risk factors, there were fewer functional risk factors identified, suggesting that there is limited availability of these data sources for use within quality measurement. Additionally, there is a lack of a generally accepted approach to defining functional status, which may contribute to a paucity of these factors. With respect to methods, regression analyses were largely used for testing. However, additional statistical approaches were applied to further assess the contribution of social and/or functional risk factors to the risk model fit.

Within the federal and state programs reviewed, similar social and functional status-related factors were identified, compared to the literature and measures reviewed. There are a variety of approaches to account for social and functional risk factors and are often used in combination, depending on the purpose, care setting, and data availability of the particular program. As measures continue to be used within these quality improvement mechanisms, additional transparency of analytical approach within and across these programs is warranted, such as selection of quality measure sets, risk adjustment models and results, and the impact of different approaches on performance ranking among entities.

Introduction

Background

The quality measurement enterprise continues to tie payment to quality of care, generally known as value-based purchasing (VBP). For VBP to be successful, patients need accurate and reliable information on provider performance (e.g., clinicians, health plans, health systems/hospitals) to make informed decisions. In addition, providers need comprehensive, reliable, and timely information to make quality care decisions that result in improved outcomes for patients while being held accountable for those outcomes in a fair and comparable manner. To level the playing field, risk adjustment methods have been applied to many measures, but not all, and not in a standardized manner across measures and programs.

Risk adjustment refers to statistical methods to control or account for patient-, community-, health plan-, or facility-level risk factors when computing outcome performance measures and resource use measures.[1] Risk adjusting measures to account for differences in patient health status and clinical factors (e.g., comorbidities, severity of illness) that are present at the start of care has been widely accepted and implemented.[2,3] However, the increased use of outcome and resource use measures in payment models and public reporting programs has raised concerns regarding the adequacy and fairness of the risk adjustment methodologies used in these measures, especially as it relates to functional status-related factors such as the ability to perform activities of daily living (e.g., eating, bathing, dressing, toileting)[4–6] and social risk factors such as income, education, social support, neighborhood deprivation, and rurality.[7,8] Functional risk factors are important to examine since they may confound the relationship between social risk, quality outcomes, and resource use. Measure developers/stewards and program implementors have long expressed a need for technical guidance on developing and testing social and/or clinical risk adjustment models, and the appropriateness of a standardized risk adjustment framework.[9] Moreover, risk adjustment of functional status-related factors within quality measurement is underexplored and underutilized for comparing provider performance on health outcomes and resource use.

Project Overview

For this effort, NQF will build on several years of work developing guidance for risk adjustment model development. Prior to 2014, NQF's guidance prohibited the inclusion of social risk factors in the risk adjustment models of measures submitted for NQF review and endorsement due to concerns of masking disparities in care.[10] NQF convened a Risk Adjustment Expert Panel in 2014, which recommended allowing risk adjustment when there is a conceptual rationale and empirical relationship.[10] The NQF Board of Directors implemented a trial period in 2015 during which adjustment of measures for social risk factors was no longer prohibited.[11] At the conclusion of this trial period in 2017, the NQF Committees and measure developers reiterated the importance of addressing all factors (both clinical and social) that can influence the result and validity of a performance measure in truly reflecting care quality and resource use.[9] These efforts have demonstrated that social risk adjustment may be feasible and appropriate, but it remains challenging for many measure developers. The increased availability of electronic data sources may be promising for addressing the issue of data availability but not for heterogeneity of social and functional risk data and modeling approaches. This suggests that exploration of electronic data sources to support functional and social risk adjustment may be a critical next step.

NQF seeks to advance measurement science in this important area by developing technical guidance for measure developers that includes emerging best practices for functional and social risk factor adjustment in measure development. The technical guidance will be informed by an environmental scan of data sources used for risk adjustment, functional or social risk factors available for testing, and approaches to conceptual and statistical methods for risk adjustment. To accomplish these goals, NQF, with funding from the Centers for Medicare & Medicaid Services (CMS), convened a multistakeholder Technical Expert Panel (TEP; Appendix A) in fall 2020 to provide input and guidance on the current state of risk adjustment for social and functional status in measurement, emerging best practices for social and functional status-related risk adjustment, the appropriateness of a standard risk adjustment framework, and the development of technical guidance for measure developers.

During the first phase of this effort, the TEP provided guidance on an environmental scan. The scan considered the use of functional, clinical, and social risk factors in measurement, and the availability and scientific acceptability of any standardized risk adjustment frameworks. In particular, NQF leveraged the experience and expertise of TEP members to identify and assess the current state of datasets used for risk adjustment, functional, clinical, or social risk factors available for testing, approaches to conceptual and statistical methods, and approaches to interpretation and decisions to include or not include functional and/or social risk factors.

Results of the environmental scan will be used to produce technical guidance for measure developers on the process of developing a risk adjustment models that consider functional and social risk factors for outcome and resource use measures, and the appropriateness of a standard risk adjustment framework. Based on the TEP's input, NQF will develop a step-by-step approach to developing a risk adjustment model that considers functional and/or social risk that is aligned with NQF measure evaluation criteria.

Key Terms and Definitions

In this report, the following key terms are used and are also included in the glossary in Appendix B:

- **Clinical adjustment** refers to adjustment for only those physiological and psychiatric attributes, which at certain levels may be associated with an increased risk of certain diseases or death.[10]
- **Functional status** is variously defined in the health field. Generally, functional status refers to an attribute that assesses how a health condition has impacted an individual's body function, body structures, and ability to participate in activities and complete basic daily tasks.[12] Functional status covers both the individual carrying out activities of daily living and the individual participating in life situations and society.[13] This includes 1) basic physical and cognitive activities such as walking or reaching, focusing attention, and communicating, as well as the routine activities of daily living, including eating, bathing, dressing, transferring, and toileting; and 2) life situations such as school or play for children and, for adults, work outside the home or maintaining a household. Furthermore, functional limitations occur when a person's capacity to carry out such activities or performance of such activities is compromised due to a health condition or injury and is not compensated by environmental factors (including physical, social, and attitudinal factors). Functional status encompasses the whole person and is affected by physical, developmental, behavioral, emotional, social, and environmental conditions.
- **Outcome** will be used broadly include the following types of outcomes relevant to performance measurement: health outcomes (e.g., mortality), intermediate clinical outcome (e.g., BP < 140/90), and economic outcomes of cost and resource use.

- **Risk adjustment** refers to statistical methods to control or account for patient-, facility-, and/or community-level factors when computing performance measure scores; methods include modeling techniques, indirect standardization, or direct standardization. These methods can be used to produce a ratio of observed-to-expected, a risk-adjusted rate, or another estimate of performance.
- **Social risk factors** are the social conditions associated with health (e.g., age, race, ethnicity, cultural context, gender, social relationships, residential and community context), including socioeconomic status (e.g., income, education, occupation).[14]
- **Social or functional status-related risk adjustment** refers to statistical adjustment for sociodemographic and/or function status-related variables.

Environmental Scan

The environmental scan was conducted using three interrelated approaches. First, a literature review was conducted to identify how risk-adjustment model development has considered social or functional risk adjustment. Second, and closely related, an examination of risk-adjustment methods used in an illustrative sample of performance measures submitted to the NQF for endorsement was conducted. Finally, federal and non-federal value-based performance measurement programs were evaluated to understand how social and/or functional risk was considered. Each of these approaches is outlined below.

Literature Review

Methods

A PubMed search was conducted of available literature published in English from the last six (6) years (since NQF's 2014 report on *Risk Adjustment for Socioeconomic Status or Other Sociodemographic Factors*) to identify studies reporting risk adjustment model development that consider social risk and/or functional status-related risk factors within quality performance measurement. Search terms included a series of terms identified through PubMed's Medical Subject Headings (MeSH), which is the National Library of Medicine-controlled vocabulary thesaurus used for indexing articles for PubMed.[15] The following MeSH terms were used for the PubMed search: "outcome and process assessment, health care"; "quality indicators, health care"; "quality of health care"; and "risk adjustment".

A reference review was also conducted of NQF's 2014 report on *Risk Adjustment for Socioeconomic Status or Other Sociodemographic Factors*[9] and the 2016 and 2020 Assistant Secretary for Planning and Evaluation (ASPE) reports.[16,17] A forward search of these reports was performed using Google Scholar to identify additional relevant articles. Lastly, NQF consulted experts in the field, including the Technical Expert Panel, to identify additional literature for inclusion.

Studies were screened for relevance based on the following inclusion and exclusion criteria (Figure 1):

Inclusion Criteria:

- Literature focused on U.S. healthcare system, and
- Literature that included empirical testing, and
- Literature focused on risk adjustment of social and/or functional status-related risk factors within the context of quality performance measurement, and
- Literature focused on risk adjustment guidance within quality performance measurement

Exclusion Criteria:

- Literature published prior to 2014 (this was established to include only those studies since NQF's 2014 report on *Risk Adjustment for Socioeconomic Status or Other Sociodemographic Factors*), or
- Literature not focusing on or not inclusive of the U.S. healthcare system, or
- Literature focused on approaches to risk adjustment modeling not within the context of quality performance measurement, or

- Literature not focusing on or inclusive of social and/or functional-status related risk adjustment within quality performance measurement, or
- Opinion papers, blogs, comments, or
- Literature that does not include empirical testing

The environmental scan prioritized outcome, and cost and resource use measures. Studies focusing solely on patient experience as an outcome were not included, as social risk adjustment has been generally accepted in this area of measurement.[18]

After screening for relevance, NQF abstracted data from included studies to capture the following:

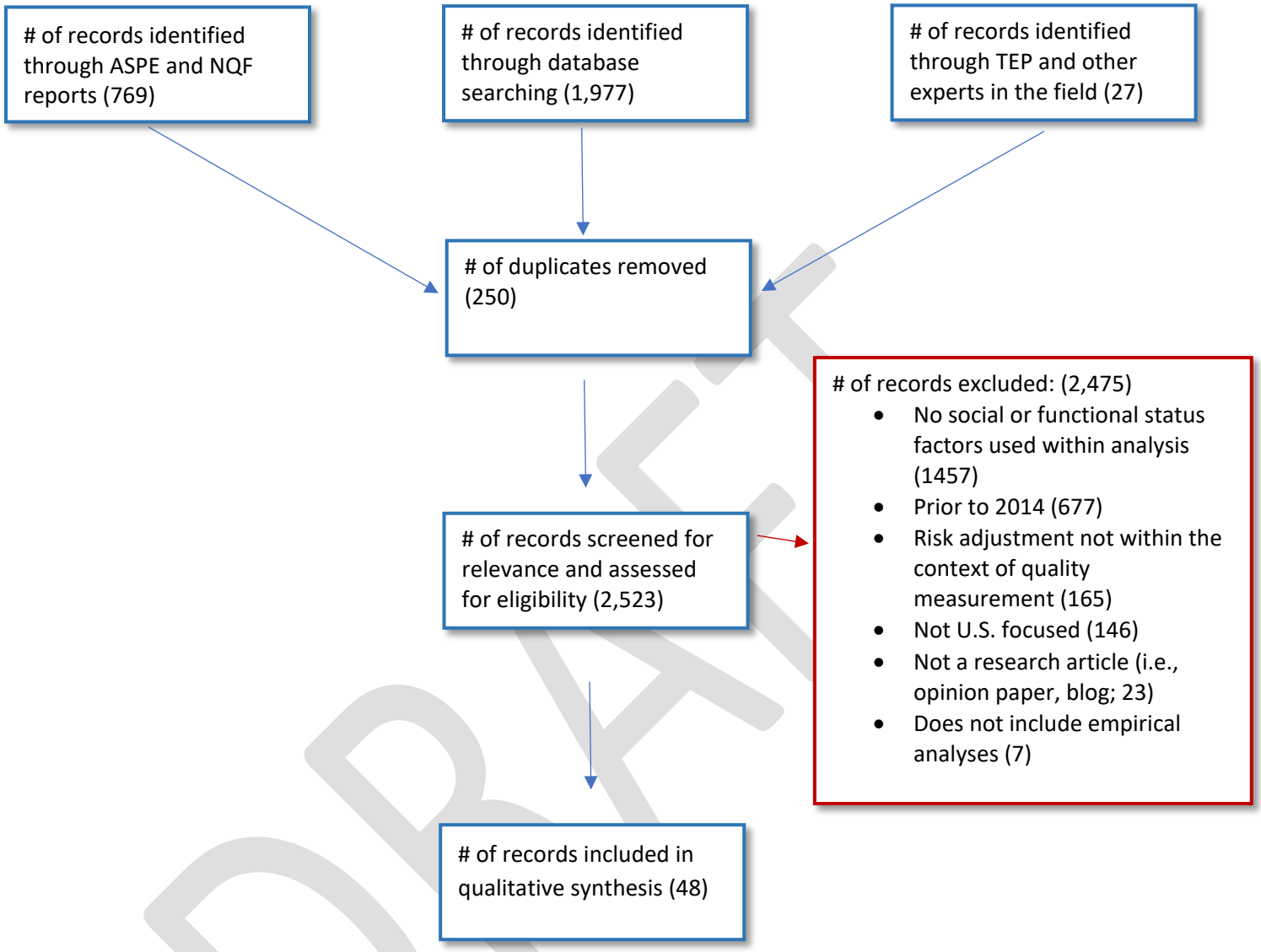
1. Datasets used: the datasets used for risk adjustment and measure specifications

2. Functional or social risk factors available for testing: the functional or social risk factors available for use in measure specification and testing

3. Approaches to conceptual and statistical methods: the conceptual and statistical methods and criteria to select patient factors (functional risk or social risk factors), as well as community-, plan-, and facility-level factors

4. Approaches to inclusion of functional and social risk factors: the myriad of approaches to examining inclusion of risk factors, including but not limited to prevalence of the factor across measured entities, empirical association with the outcome, contribution of unique variation in the outcome, and assessment of between-unit effects and within-unit effects

Figure 1. Literature Flow Diagram



Measure Review

Methods

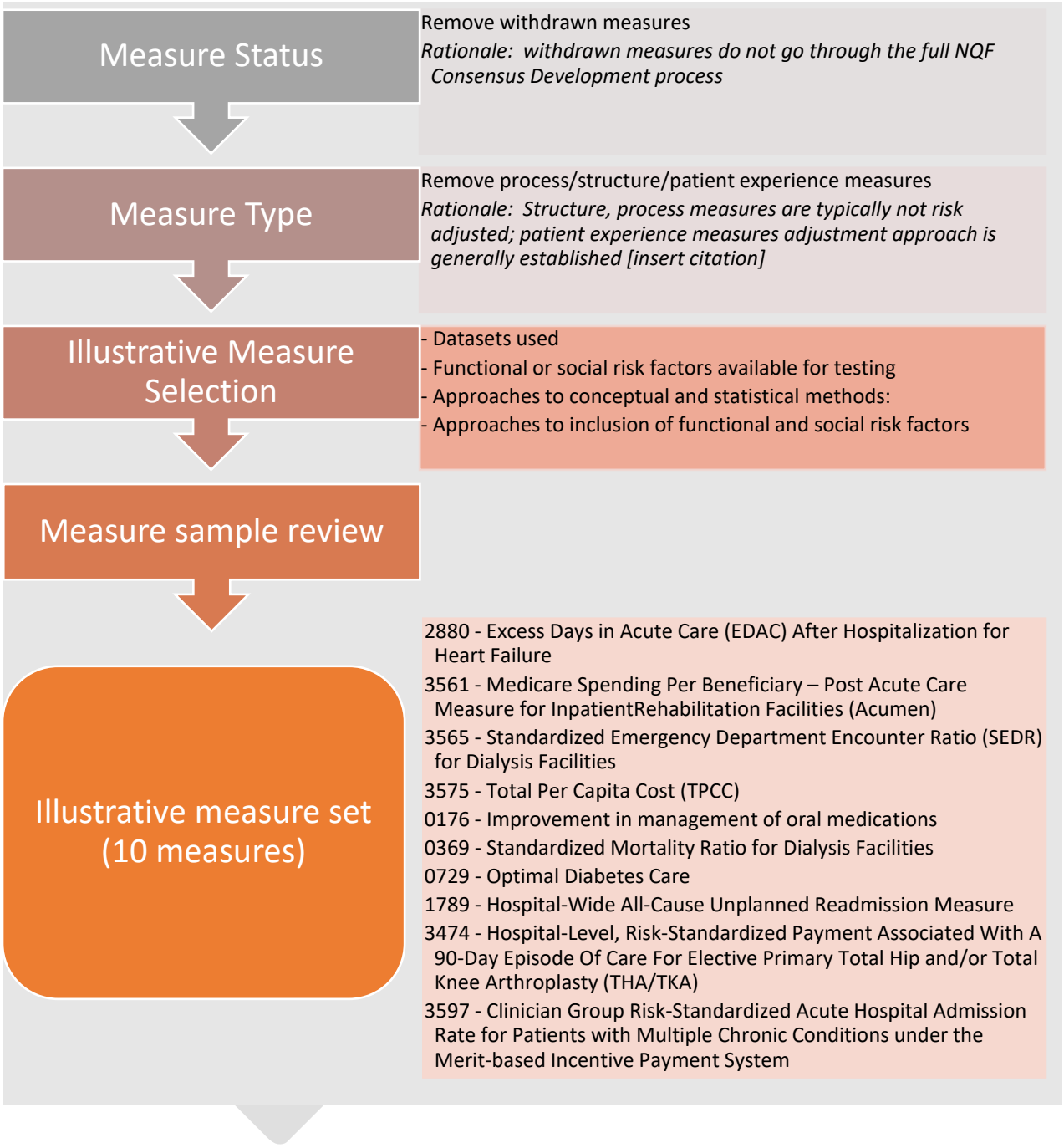
Measures from the NQF-endorsed measure portfolios and candidate measures submitted for NQF endorsement were reviewed for potential inclusion as illustrative measures for the environmental scan. NQF prioritized illustrative measures with novel or robust approaches to measure testing in NQF measure submissions forms (also called testing attachments) previously submitted to the Consensus Development Process (CDP) projects to identify 10 measures that showcase datasets used, functional or social risk factors available for testing, approaches to conceptual and statistical methods, and a range of care settings and target populations.

In all, 10 illustrative measures were selected for presentation and analysis of these considerations. Regarding approaches to conceptual and statistical methods, NQF examined the "ordering" of risk factor inclusion (e.g., are social risk factors added before or after all clinical factors). NQF also examined the relationship between functional risk adjustment and social risk adjustment by measure type and intended use.

Utilizing the measure selection logic presented below, NQF identified all measures across multiple CDP projects under evaluation in the 2017-2021 NQF social risk trial. First, all measures withdrawn from NQF endorsement consideration were removed from consideration in this project. All process and structure measures were removed from consideration for inclusion in this project since these measures should only be adjusted in particular circumstances.[10] Namely, the process or structure is indicated for all patients within the denominator and adjustment is rarely required.[19] Patient experience measures were removed from consideration in this project, as risk adjustment in patient experience measures is generally well established.

The environmental scan prioritized outcome and cost and efficiency measures for which a conceptual rationale for adjustment was demonstrated. NQF conducted a preliminary review of this subset of submitted testing attachments to identify an illustrative set of 10 measures. Figure 2 below illustrates this process of measure identification.

Figure 2. Measure Flow Diagram



Program Review

Methods

The environmental scan examined various federal and non-federal value-based payment and/or public reporting programs that include quality measures. The scan prioritized programs that adjust or stratify performance measures for social and functional risk factors at the program level. The programs selected may also include adjustment at the individual measure level. The program review did not intend to comprehensively examine all the federal and non-federal payment or public reporting programs, but rather to identify illustrative examples in which social and functional risk factors are accounted for at the program level. Since the program review is aimed at searching for illustrative examples, we used a "snowball" search strategy (i.e., emerging as the study unfolded) as this method has been found especially powerful for identifying high quality sources for special topics.[20] The search was initiated with a review of federal programs included in the 2016 National Academies of Sciences Engineering and Medicine report, *Accounting for Social Risk Factors in Medicare Payment: Identifying Social Risk Factors* [14], 2020 NQF's report of *Measure Sets and Measurement Systems: Multistakeholder Guidance for Designing and Evaluation*[21], and 2020 ASPE's *Report to Congress*.[22] Next, NQF sought input from the federal liaisons and the TEP for additional federal, state and private programs exemplifying novel approaches to analyze.

Utilizing the program selection and data extraction flowchart (Figure 3) presented below, NQF identified illustrative programs that meet the following inclusion criteria:

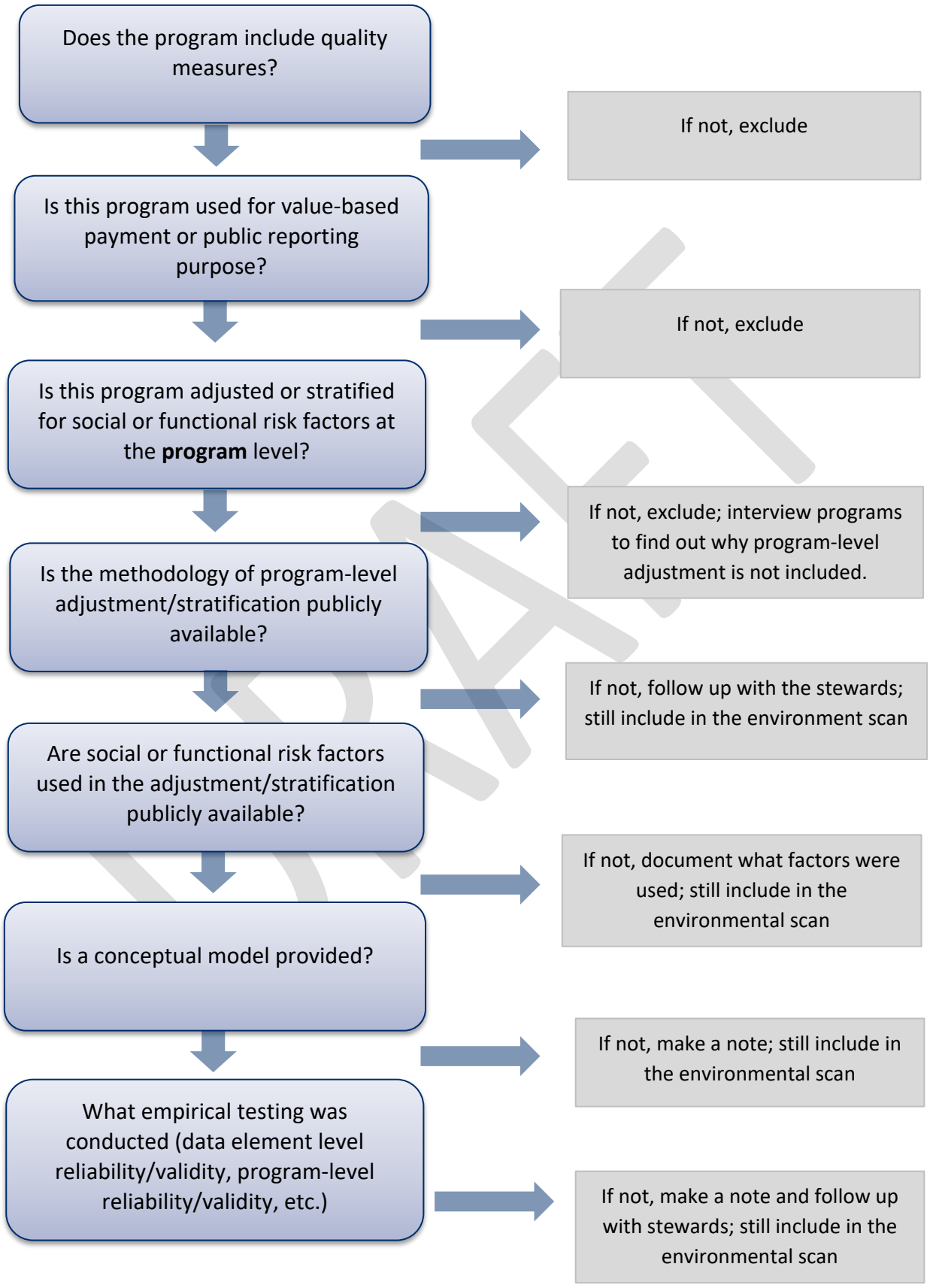
Inclusion criteria:

- The program is used for value-based payment and/or public reporting
- The program includes quality measures
- The program adjusts or stratifies for social and/or functional risk factors at the program level and, in some cases, includes adjustment at the individual measure level

Exclusion criteria:

- The program is only used for internal quality improvement
- The program does not include quality measures
- The program does not adjust for social and/or functional risk factors at the program level

Figure 3. Flowchart and Data for Program Review



Environmental Scan Findings

Literature Review Findings

A total of 2,773 articles were identified, of which, 769 were from the ASPE and NQF reports, 1,977 were identified from PubMed, and 27 more through TEP members or other experts in the field (Figure 1). After removing 250 duplicates, 2,523 articles were screened for relevance and assessed for eligibility. An additional 2,475 records were excluded based on the inclusion and exclusion criteria. Specifically, 1,457 records did not include social or functional status factor within the risk adjustment analysis, 677 records were published prior to 2014, 165 records focused on risk adjustment that was not within the context of healthcare quality measurement, 146 records did not focus on the U.S. healthcare system, 23 records were not research articles, and seven records did not include empirical analysis. This yielded 48 articles for the qualitative analysis.

It should be noted that the findings from the literature may be limited in their generalizability because of the data sources and subsequent social and/or functional status-related risk factors that were available. Several studies used non-public data (i.e., health system data, or data from a commercial entity such as a health plan) and/or narrowly focused data (i.e., within a specific health system). Therefore, some of the risk factors may not be widely available or feasible to use beyond the scope of the study or outside the healthcare entity upon which the study was based. Information regarding the data sources used, and their accessibility and use, can be found in Appendix C.

The literature included in the environmental scan covered several topic areas (Table 1). Twenty-seven studies focused on cost/resource use measures, such as admissions and/or readmissions to the hospital and episode-based payments. Mortality and patient survival rates were included in 18 studies. The remaining studies included patient safety indicators (e.g., infection rates), disease control metrics (e.g., blood sugar control), and disease severity/complications (e.g., stroke).

With respect to the level of analyses, provider accountability measured, thirty-three studies focused on the hospital/facility level; four studies examined clinician/clinician group level performance; two studies investigated accountable care organizations or patient-centered medical homes; and two studies examined health plans. Lastly, these studies all included analysis of risk adjustment for social and/or functional status-related risk factors. Finally, 43 studies included an analysis of social risk factors and 13 included functional risk factors.

Data Sources Used (Table 2)

In the studies we examined, a range of data sources were used to calculate the measure and for social and/or function status risk factor analyses (Table 2). The most frequently used data sources were the Medicare claims (Part A and Part B) as well as registry data and electronic health records, all of which included variables used for the measure calculation and for the social risk factor analysis. Non-public sources, such as data from health systems or commercial entities, were also frequently used. For the social risk factor analysis alone, the American Community Survey was a commonly used data source. With respect to functional status-related risk factors, a range of data sources were used, including the Medicare Current Beneficiary Survey, the Health and Retirement Study-Medicare Linked Dataset, and registry data.

Approaches to empirical analyses of social/functional status-related risk factors (Table 3)

This scan identified a number of analytical approaches for social risk factor analysis, including univariate and bivariate testing, the incremental effect of social and/or functional status-related risk factors in a multivariable model, net reclassification index analyses, an assessment of the contribution of these risk factors to the risk model fit, and an assessment of correlation of the adjusted score/unadjusted score (Table 3).

Several studies conducted univariate or bivariate testing to examine the contribution of social risk and/or functional status-related risk factors in predicting the measure's outcome. Risk factors that were not deemed statistically significant were removed from the risk adjustment model. Multivariable analyses were frequently used to examine the contribution of the social/functional status-related risk factors to the outcome measure.

Three studies included a net reclassification improvement (NRI) index. The NRI is an index measure of how well a new model reclassifies subjects compared with an old model (correct versus incorrect changes in prediction for cases and controls separately).[23–25] In the current context, NRI is used to identify whether the addition of social and/or functional risk factors results in models that more accurately classify outcomes. The goal of the NRI is to compare the shifts in reclassified categories by outcome and results from the addition of the social and/or functional status-related risk and the clinical covariates to the model. A higher NRI index indicates greater improvement in risk discrimination and better reclassification. Three studies included a net reclassification improvement (NRI) index.

Lastly, several studies examined the contribution of social/functional status-related risk factors to the model fit by evaluating whether these factors improved model performance (e.g., the model adjusted r-squared value, calibration, c-statistic, Brier score). Eleven of the studies assessed the correlation of the unadjusted outcome with social/functional status-related risk factors being included in the risk model.

Social and functional risk factors identified (Table 4)

We found a wide range of social and functional status-related risk factors included in the literature (Table 4), including age, gender, race, ethnicity, income, insurance status (e.g., Medicare/Medicaid dual eligibility, low-income subsidy), socioeconomic status (SES), and community health indices which based on beneficiary zip code or place of residence. An example of these indices includes the Agency for Healthcare Research and Quality (AHRQ) SES Index,[26] which summarizes area-level measures of employment, income, education, and housing. Each of the index components is available at the census block level, which was then linked to patient's residence using nine-digit ZIP code. Another index used was the Distressed Communities Index (DCI).[27] The DCI is based on seven metrics: no high school degree, housing vacancy rate, adults not working, poverty rate, median income ratio, change in employment, and change in business establishments, which can be obtained from the American Communities Survey.

Most of the included studies used patient-level information for social risk. Although not as frequent, community-level information (e.g., averages, percentages, proportions) was also used and included representative: age, race, gender, education level, food access, air pollution, access to exercise opportunities (i.e., gyms, fitness centers in the area), income, including poverty rates, percentage of unemployment, and the SES/community health indices mentioned previously.

With respect to function, risk factors included activities of daily living (ADLs), instrumental activities of daily living (IADLs), ambulatory function (e.g., gait speed, fall risk), cognitive impairment, and overall functional status, which was evaluated based on a person's dependence to accomplish tasks. One study used a propriety risk score, which included a functional status assessment. Although not detailed in the table below, survey data were largely used to identify functional status-related factors (e.g., the Medicare Current Beneficiary Survey provides survey data on ADLs and IADLs). Lastly, patient-level information was used for all the functional status-related risk factors.

Measure Review Findings

The measures included in the environmental scan cover several topic areas and levels of analyses. Table 1 below details the summary findings. Four measures were extracted from the All-Cause Admissions and Readmissions portfolio, three from the Cost and Efficiency portfolio, and one each from each of the following: the Geriatrics and Palliative Care, Primary Care and Chronic Illness (PCCI), and Renal portfolios. Most of these measures are outcomes, but three are cost measures, and one (from the PCCI portfolio) is a composite measure.

These measures also cover a range of levels of analyses. Three measures are at the clinician/clinician group level, one is at the accountable care organization (ACO) level, and seven measures are at the hospital level. One measure, NQF #1789, is specified and analyzed for both Facility/Hospital and ACO levels of analysis. Similarly, the measures have a range of intended uses. Seven measures are intended for public reporting and five measures are intended for payment applications. Finally, all 10 ten measures were selected because social and/or functional status-related risk factors were analyzed for inclusion in a risk adjustment model; of these, only four included an analysis of functional risk factors.

Data sources used (Table 2)

Given the measure focus areas and target populations, the performance measures identified in the scan used data from eight data sources frequently used in the quality measurement field (Table 2). These data sources include Medicare Part A claims, Medicare Part B claims, Medicare Enrollment Database, Chronic Condition Data Warehouse (includes Medicare Part A and Part B claims data, Medicare Enrollment Data, and Part D Prescription Drug Event data), Provider of Services File, American Community Survey, Common Medicare Environment database, and USDA Rural Urban Continuum Codes. The data sources are used for social risk factor and functional risk factor analyses and calculating the final measure result. The most frequently used data source overall was the Medicare Enrollment Database (MED), which includes variables used in final measure calculation and in social risk factor analysis. The Chronic Condition Data Warehouse (CCW) was used for functional risk factor analysis, and the American Community Survey was used for social risk factor analysis.

Other datasets used included Home-Health OASIS-C2, Dialysis Facility Compare, and registry data from CROWNWeb, the Renal Management Information System, and the Scientific Registry of Transplant Recipients.

Approaches to empirical analyses of social risk factors (Table 3)

This measure review identified a number of analytical approaches for social and/or functional status-related risk factor analysis, including: an examination of variation in prevalence of the social risk factor across measured entities; bivariate testing; the incremental effect of risk factors in a multivariable

model; decomposition analysis; the contribution of social risk factors to the risk model fit; and an assessment of correlation of the social/functional status risk score with the unadjusted score (Table 3).

Four measures conducted bivariate testing to examine the contribution of social risk factors in predicting the measure's outcome. In one case, clinicians reviewed the strength and direction of the relationship between the social risk factor and the measure outcome, and risk factors identified in bivariate testing that were not deemed clinically plausible were removed from the risk model. Of the four measures that conducted bivariate testing, two did additional multivariate testing of the contribution of the social risk factors to the measure outcome. In all, three measures used multivariate testing to examine the strength and significance of the social risk factor variables in the context of a multivariate model that included claims-based clinical variables.

Two measures in the study sample used decomposition analysis to examine the impact of social risk factors. The goal of this analysis was to disentangle the patient-level and hospital-level social risk factors in the risk adjustment models. If both the patient-level and hospital-level effects are significant, that indicates that both are associated with increased risk of the outcome being measured. In this case, if social risk factors are used in the final risk adjustment model to adjust for patient-level differences, some of the variation in risk of the measured outcome between hospitals would be adjusted for, potentially obscuring a quality signal.

One measure analyzed referral patterns by level of analysis (e.g., hospital) characteristics. To further understand the effect of social risk factors on the measured outcome of payments following hip/knee surgery, the measure stratified testing results by 1) dual and non-dual eligible patients and 2) low SES and non-low SES patients using the AHRQ SES Index score. Further, to understand whether this association between social risk and payments was driven by a patient- or hospital-level effect, the developer examined referral patterns and observed payments for dual eligible and non-dual eligible patients among hospitals with a high overall proportion of dual eligible patients and hospitals with a low overall proportion of dual eligible patients.[28] These analyses helped to reveal whether patterns of use of post-acute care settings and payments associated with that care were driven mostly by the patient's dual eligible status or the fact that they received care at a hospital that cares for a large proportion of dual eligible patients. These analyses were replicated for hospitals with high and low proportions of low SES patients using the AHRQ SES index.

All 10 measures assessed the contribution of social risk factors to the risk model fit by evaluating whether the social risk factor increased the risk adjustment model performance statistics (e.g., model calibration, adjusted r-squared value and/or c-statistic). Further, five of the measures were used to conduct an assessment of the correlation of the measure score with and without social risk factors included in the risk model using Pearson and/or Spearman correlation tests. Assessments of the contribution of social risk factor adjustment to model fit and correlation of measure scores were conducted using measure scores with clinical factors already included in the risk model.

Social and functional risk factors identified (Table 4)

The social risk factors that were identified in the measure review include race, ethnicity, Medicare/Medicaid dual eligibility, and socioeconomic status index based on beneficiary zip code (Table 4). An example of this index includes the Agency for Healthcare Research and Quality (AHRQ) SES Index, which summarizes area-level measures of employment, income, education, and housing. Each of the

index components is available at the census block level, which were linked to patient's residence using nine 9-digit ZIP code. Of note, variables were identified as proxies for the risk factors identified above (e.g., payment source data from OASIS-C2 database were used as a proxy for dual eligibility).

Of the 10 measures in the environmental scan, five measures used Medicare/Medicaid dual eligibility as a risk factor in their analysis, six measures used race and/or ethnicity as a potential factor, six used urbanicity (as based on ZIP code), and eight measures used AHRQ SES. All but one measure used patient-level information to collect these social risk factors. One measure used community-level data to collect information on urbanicity and SES factors. No measures used facility-level data for adjustment.

Of the four measures identified in the scan that assessed functional risk, risk factors included psychiatric disorders, functional disabilities, and conditions that limited activities of daily living such as ambulation and grooming. Additionally, although not reflected in the table below, claims data were used to identify proxies for frailty as a functional risk factor (e.g., a billing claim for home oxygen or a hospital bed lift).

Program Review Findings

Performance measures have been used in numerous ways to drive quality improvement, including public reporting, pay-for-performance, and value-based purchasing programs. Measure results are usually reported at the individual performance measure level and typically then aggregated up to a single rating (e.g., stars) for the accountable entity (e.g., plans or hospitals). An aggregated performance score for each accountable entity enables consumers, payers, and purchasers to compare performance across entities and thus supports decision-making in selecting providers.

Many of these publicly reported measures are then also used as basis for quality-based financial incentives. As the 2016 National Academies of Sciences Engineering and Medicine report [16] describes different categories of methods to account for social risk factors may achieve different policy goals and may have different consequences, some of which may be unintended and adverse. Often a program may use a combination of those approaches, depending on the purpose, care setting, and data availability of the particular program. In the programs reviewed, risk adjustment of social and functional status-related factors often occurred at the individual measure-level (such as CAHPS measures) or approximated at a higher level (such as measure stars). One program stratified program performance by the reporting entity's social risk factors. We include the following figures to illustrate how various programs incorporate risk adjustment for social and functional status in their evaluations (Figures 4-6).

Three general risk adjustment approaches for social and functional status in the design of a program emerged through the program review:

1. Program performance is adjusted for mean within-entity differences

The Medicare Advantage (MA) and Prescription Drug Plan's (PDP) Star Ratings consists of over 50 measures. Among those measures, Healthcare Effectiveness Data and Information Set (HEDIS) measures are not adjusted for individual social or functional risk factors. The Categorical Adjustment Index (CAI) was introduced in 2017 to adjust for the average within-contract disparity in performance among beneficiaries who receive a low-income subsidy (LIS), are Medicare/Medicaid dual eligible (DE), and/or are disabled. The MA and CAI approximates the effect of within-entity adjustment of HEDIS measures by

translating that effect into stars. CMS updates the CAI values by examining the within-contract differences in performance between LIS/DE and non-LIS/DE beneficiaries for measures that are not adjusted. To calculate the CAI, CMS determines adjusted measure scores for all candidate measures that approximate case-mix adjustment using a beneficiary-level logistic regression model with contract-fixed effects and beneficiary-level indicators of LIS/DE and disability status. Based on the results of those statistical models, this set of measures is then adjusted at the measure-star level and rolled up to the overall and summary stars (Figure 4). For the 2021 Star Ratings, CAI adjustment included 19 Part C and Part D measures.[29]

2. Program performance is stratified by reporting entity characteristics

For example, the CMS Hospital Readmissions Reduction Program (HRRP) was established to reduce payments to Inpatient Prospective Payment System (IPPS) hospitals for excess readmissions. Hospitals are stratified into five peer groups, or quintiles, based on the proportion of patients that are dual eligible for both Medicare and full-benefit Medicaid. Hospital performance is then assessed relative to the performance of hospitals within the same peer group along six quality measures. The median excess readmission ratios (ERRs) of hospitals within the peer group is used as the threshold to assess hospital performance on each measure.[30]

3. Performance measures included in the program are directly adjusted for social and/or functional risk factors at the individual measure level without additional adjustment or stratification in the program in which they are deployed. (Many performance measures are adjusted for clinical risk factors only; they are not listed below.) A few examples with varying characteristics include:

    - MA and Part D Star Ratings include MA and PDP CAHPS measures that are directly risk-adjusted for respondent's social risk factors that are not under the control of the health or drug plan but related to the sampled member's survey responses, such as age, education, physical and mental health status, Medicaid, LIS, proxy respondent, help with responding and Chinese version survey used. The Health Outcomes Survey (HOS) measures are also adjusted for patient-level risk factors.

    - CMS Nursing Home Compare (Five-Star Quality Ratings System) publicly display ratings for each of quality domains as well as an overall rating based on a set of quality ratings for each nursing home that participate in Medicare or Medicaid. One of the domains is MDS-based quality measures, including mobility decline, catheter, short-stay functional improvement, and short-stay pressure ulcers. They are risk adjusted for residents' demographics, clinical risks, cognitive impairment, and long-form activities of daily living (ADL) score (eating, toileting, transfer, and walking in corridor).[31]

    - CMS Home Health Program includes three sets of quality measures for public reporting (Figure 5): Medicare.gov website publishes a large set of quality measures based on the Outcome and Assessment Information Set (OASIS) assessments and Medicare claims data. Seven of these measures are then consolidated into the Quality of Patient Care (QoPC) Star Ratings. OASIS-based measures used in the star ratings are adjusted for demographics, payment sources, clinical risks, and functional status such as grooming, upper body dressing, lower body dressing, etc. OASIS also collects and adjusts for race and ethnicity through completion of the OASIS instrument. A separate Patient Survey Rating, based on HHCAHPS,

is also reported on Home Health Compare, which adjusts for social risk factors such as education, residence status (living alone), non-English survey response, in addition to age, clinical diagnoses and self-reported health and mental health status.[32]

On the payment side, the case-mix methodology for rate setting includes functional impairment level (low, medium, high) based on OASIS items, in addition to admission source and timing, clinical grouping and comorbidity.[33]

In January 2021, the CMS Innovation Center announced that the Home Health Value-Based Purchasing (HHVBP) Model has been approved for expansion. The Model's first performance year began on January 1, 2016 and the Model will end on December 31, 2022. Medicare-certified HHAs in the nine model states receive a Total Performance Score (TPS) on quality measures collected from OASIS, HHCAHPS and claims and three new measures where points are awarded for reporting data. The TPS is used to calculate an annual payment adjustment that began in CY 2018 at upward/downward of 3% and incrementally increases to 8% in CY 2022. The two claims-based measures, *Emergency Department (ED) Use without Hospitalization* (NQF #0173) and *Unplanned Acute Care Hospitalization* (NQF #0171), are both adjusted for an array of clinical and demographic risk factors, as well as disability. They carry a larger weight in the TPS calculation in the later years since the inception of HHVBP.[34,35]

- CMS's End Stage Renal Disease (ESRD) Quality Incentive Program (QIP), which reduces payments to ESRD facilities that do not meet performance standards, includes the standardized mortality ratio (SMR) as a key quality measure, which adjusts for age, gender, race (white, black, Asian/PI, Native American or other) and ethnicity (Hispanic, non-Hispanic, or unknown).[36,37] The In-Center Hemodialysis (ICH) CAHPS Survey which is part of the QIP adjusts for social risk factors such as education and speaking a language other than English in the home, in addition to age, heart disease, difficulty seeing and hearing, difficulty with activities of daily living, and self-reported health and mental health status.[38]

- New York Managed Long-Term Care Plans: New York State certifies and oversees the operation of New York State managed long-term care (MLTC) plans. New York State Department of Health (NYSDOH) has been publishing quality performance and enrollment data for MLTC plans since 2012, including a set of measures for the current plan performance (quality of life, effectiveness of care, emergency room visits, access and experience of care), a set of measures for plan performance over time (functioning and activities of daily living, quality of life and effectiveness of care), the rate of potentially avoidable hospitalizations (PAH) per 10,000 days enrolled in the plan, and a member satisfaction survey on experience of care.[39] Based on MLTC member assessments, the NYSDOH developed a functional assessment scoring system, the Nursing Facility Level of Care (NFLOC) score. The NFLOC score is comprised of 11 components that are derived from 22 items from the Uniform Assessment System for New York (UAS-NY) instrument. The items include the areas of incontinence, cognitive performance, activities of daily living (ADLs), and behavior. The NFLOC score is used as an unadjusted descriptive measure and as an adjusted over-time measure. The MLTC assessment data is also a source of functional risk factors in risk adjustment models:

- Currently, five measures for current performance and nine measures for performance over time are risk adjusted for clinical factors as well as functional status. The models may include factors such as supervision through total dependence in locomotion; unsteady gait present; walks with assistive device, uses wheelchair, or is bed bound; not independent in bathing; total dependence in ADL locomotion, hygiene, and bathing; supervision through total dependence in managing medications. The risk adjustment models are calibrated every year and only the significant covariates are kept in the models.
- Currently, Potentially Avoidable Hospitalization adjusts for clinical factors and functional status, including extensive assistance through total dependence in locomotion; unsteady gait present; walks with assistive device, uses wheelchair, or is bed bound; decline in ADL status compared to 90 days ago; supervision through total dependence in managing medications.
- Some measures from the survey of satisfaction with experience of care are adjusted for age, education, self-rated health status. Two measures are adjusted for cognition.

On the payment side, some of the adjusted quality measures are used as basis for a pay-for-performance financial incentive as those measures compare current year's measure results between plans. An additional VBP shared savings program uses the *unadjusted* quality measures to provide financial incentive if targets are met over time for a set of specific quality measures agreed between plans and provider.[40]

Social and functional risk factors identified

Among the programs reviewed, social risk factors used to adjust program-level performance include Medicare/Medicaid dual eligibility and disability. Race and ethnicity are included because such information is collected through patient portals for ESRD patients. All of these programs include patient experience with care surveys, which adjust for self-reported social factors such as education and language of survey. Regarding functional status, instruments including MDS and OASIS are used both as quality measures to assess changes in functional status and as risk adjustors for various performance measures for settings of nursing homes, home health, and long-term care.

Figure 4. Medicare Advantage and Part D Star Ratings

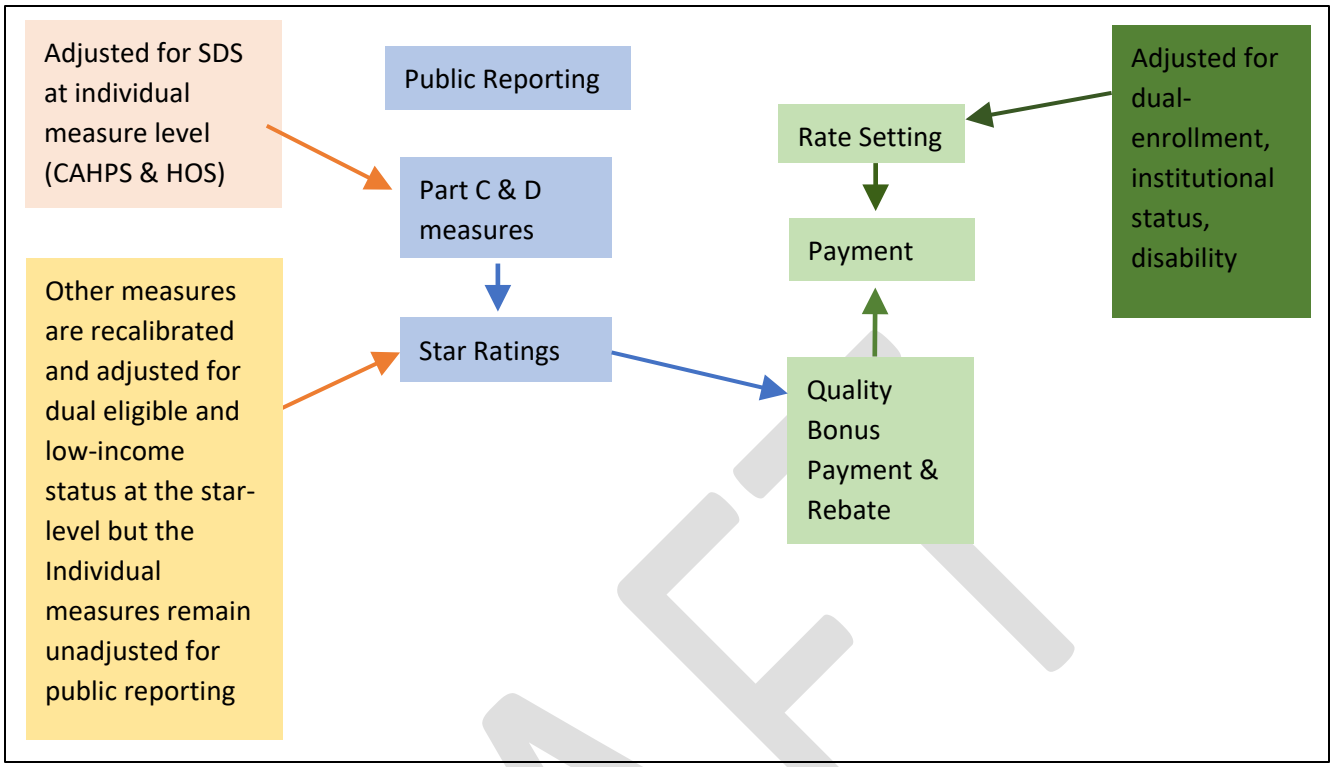


Figure 5. CMS Home Health Program

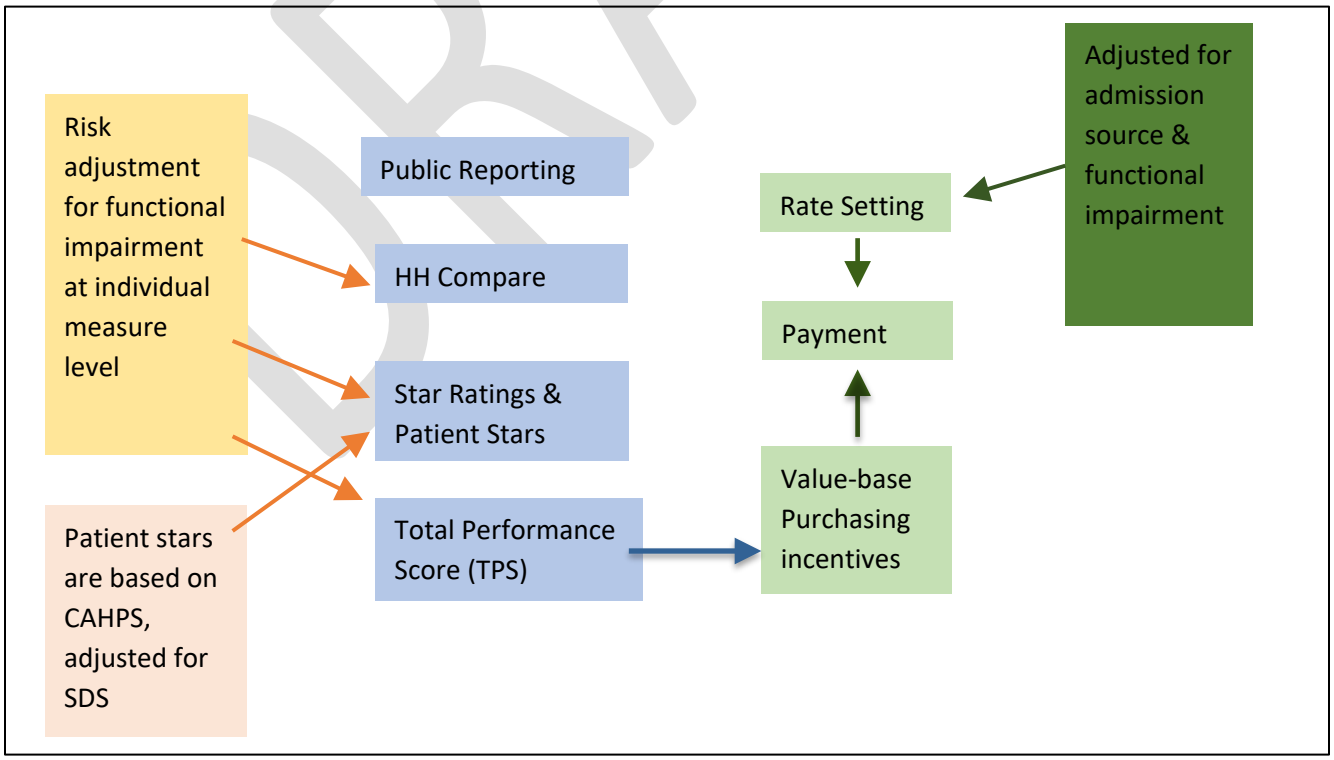
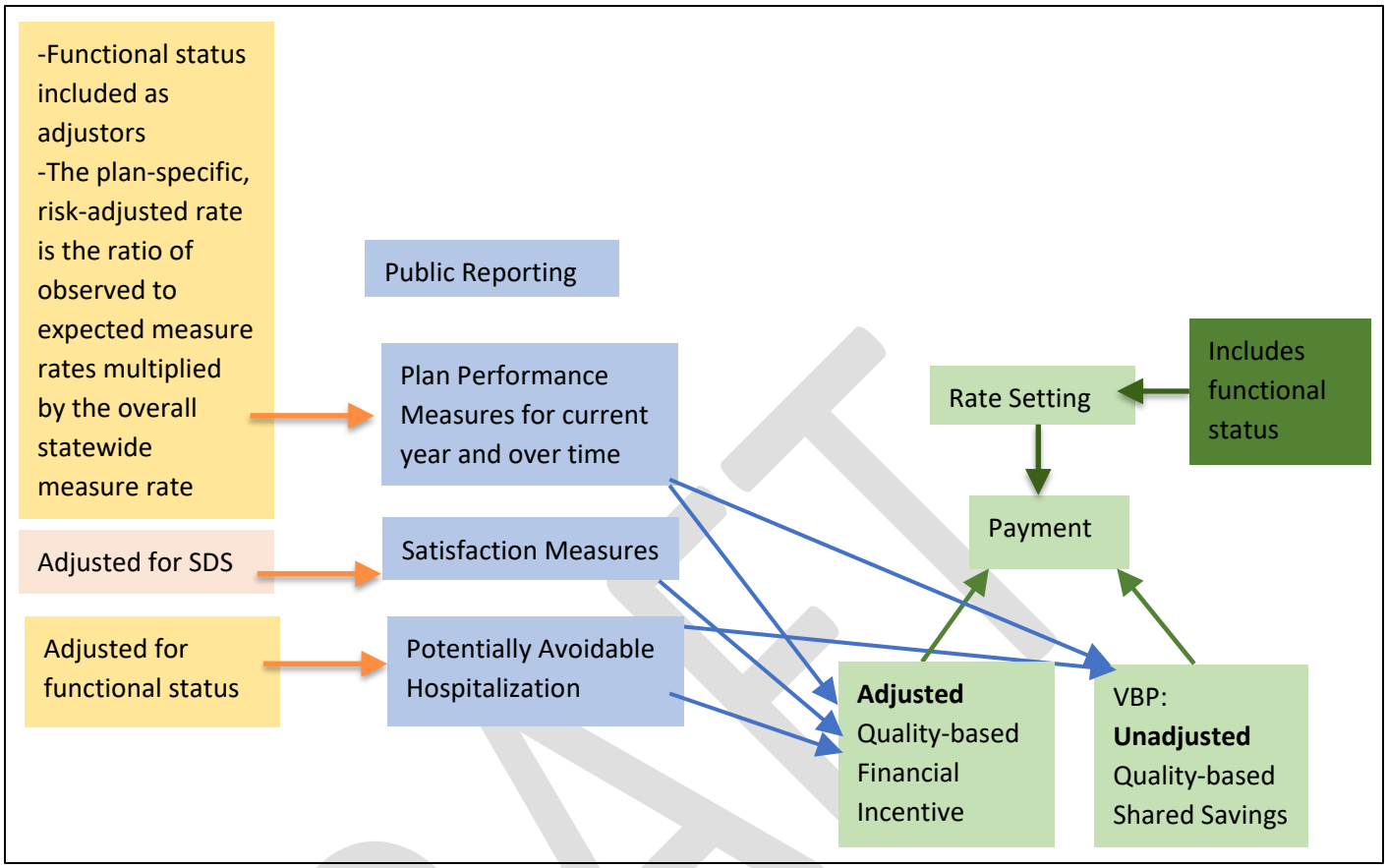


Figure 6. New York Managed Long-Term Care Plans



Note: Various functional status data elements are considered for each model each year. Only those data elements that are significant are retained in the final risk-adjusted model.

Table 1: Descriptive Characteristics of Measures Used (Counts)*

	Literature Analysis	*Measure Analysis*
**Total number (counts) of illustrative examples reviewed. The data presented below are not intended to be an exhaustive review.*		
Measurement Focus Area		
Admissions and Readmissions	-	4
Cost	-	3
Geriatrics and Palliative Care	-	1
Primary Care and Chronic Illness	-	1
Renal	-	1
Total	-	*10*
Measure Type		
Outcome measures (e.g., disease management, infection rates)	13	5
Stroke	1	-
Mortality	18	1
Cost/Resource Use measure	27	3
Composite	1	1
Total	*60*	*10*
Level of Analysis		
Clinician/Clinician Group	4	3
Accountable Care Organization (ACO)/Patient-Centered Medical Home (PCMH)	2	1
Geographic (e.g., rural-urban, community)	2	-
Hospital/Facility	38	7
Health Plan	2	0
Total	*48*	*11*[1]
Risk Factor Analysis		
Social risk factors analyzed for inclusion in risk adjustment model	43	10
Functional risk factors analyzed for inclusion in risk adjustment model	13	4
Social/functional status risk factor included in final risk adjustment model	32	9

[1] One measure, NQF #1789, is specified and analyzed for both Facility/Hospital and ACO levels of analysis.

Table 2: Data Sources Used for Measure Calculation, Risk Adjustment Analysis (Counts)*

	Literature Analysis			*Measure Analysis*		
	Measure calculation	Social risk factor analysis	Functional risk factor analysis	Measure calculation	Social risk factor analysis	Functional risk factor analysis
**Total number (counts) of illustrative examples reviewed. The data presented below are not intended to be an exhaustive review.*						
American Community Survey	-	6	1	1	5	-
Area Health Resources File	-	-	-	-	1	-
Chronic Condition Data Warehouse	-	-	-	2	-	2
Centers for Medicare & Medicaid Services Provider of Services (POS) Files	1	1	-	1	-	-
Common Medicare Environment Database	-	-	-	1	2	-
Community Health Rankings	1	2	-	-	-	-
Dialysis Facility Compare	-	-	-	1	-	-
Electronic Health Record	3	3	-	-	-	-
Inpatient Rehabilitation Facility Assessment Instrument	-	-	1	-	-	-
Health and Retirement Study-Medicare Linked Dataset	2	1	2	-	-	-
Healthcare Cost and Utilization Project (HCUP) Database	1	2	-	-	-	-
MarketScan© Commercial Claims and Encounter Data	1	1	-	-	-	-
Medical Expenditure Panel Survey (MEPS)	1	1	-	-	-	-
Medicare Current Beneficiary Survey	2	2	2	-	-	-
Medicare Enrollment Database	2	3	-	6	4	2
Medicare Part A claims	10	3	-	6	2	-
Medicare Part B claims	9	1	-	3	-	-
Medicare Part C claims	1	1	-	-	-	-
Medicare Part D claims	1	1	-	-	-	-
Medicare Public Use Files	1	1	-	-	-	-
Minimum Data Set	-	-	-	3	-	-
Non-public Sources (i.e., hospital data, commercial entity)	9	9	-	-	-	-
OASIS-C2[2]	-	-	-	1	1	1
Registry Data	11	11	3	2	2	2
Rural Urban Continuum Codes	-	-	-	-	1	-
United States Census Data	-	3	-	-	-	-

[2] Outcome and Assessment Information Set

	Literature Analysis			*Measure Analysis*		
	Measure calculation	Social risk factor analysis	Functional risk factor analysis	Measure calculation	Social risk factor analysis	Functional risk factor analysis
United States Dialysis Facility Reports	1	1	-	-	-	-
Veterans' Affairs Administrative Claims/Medical Record Data	3	3	1	-	-	-

Table 3: Approaches to Empirical Analysis of Social/Functional Status-Related Risk Factors (Counts)*

	Literature Analysis	*Measure Analysis*
**Total number (counts) of illustrative examples reviewed. The data presented below are not intended to be an exhaustive review.*		
Univariate testing	5	-
Bivariate testing	3	4
Multivariate testing	30	3
Negative binomial regression	1	-
Stepwise regression	7	-
Hierarchical modeling	10	2
Decomposition analysis	-	2
Net reclassification improvement (NRI) index	3	-
Assessment of contribution to risk model fit (e.g., r-squared, c-statistic)	23	10
Assessment of correlation of social/functional status risk score/unadjusted score	11	5

Table 4: Social and/or Functional Status-Related Factors (Counts)*

	Literature Analysis			*Measure Analysis*		
	Patient-level	Facility-level	Community-level	Patient-level	Facility-level	Community-level
**Total number (counts) of illustrative examples reviewed. The data presented below are not intended to be an exhaustive review.*						
Social Factors						
Demographics						
Age	28	-	3	2	-	-
Race/ethnicity	26	-	3	6	-	-
Gender	28	-	-	3	-	-
Language	1	-	1	-	-	-
Education						
Level of education	1	-	6	-	-	-
Employment Status						
Unemployment	-	-	3	1	-	-
Environmental Factors						
Air pollution	-	-	1	-	-	-
Exercise opportunities	-	-	1	-	-	-
Food access/insecurity	-	-	2	-	-	-
Safety net hospital	-	1	-	-	-	-
Time to seek medical treatment	1	-	-	-	-	-
Geographic Location						
Street address	1	-	-	-	-	-
Urban-rural classification	3	-	-	6	-	1
US Census tracts	3	-	1	-	-	-
Zip-code	1	-	1	3	-	-
Housing/Living Situation						
Living at home	-	-	1	-	-	-
Nursing home residence	1	-	-	-	-	-
Income						
Annual personal income	2	-	-	-	-	-
Household income	1	-	4	-	-	-
Income proxy (e.g., copayment determined by income)	2	-	-	-	-	-
Net worth/wealth	2	-	-	-	-	-
Poverty	1	-	5	-	-	-

Insurance Status						
Dual eligibility (Medicare and Medicaid)	6	-	-	5	-	-
Medicare/Medicaid eligibility	4	-	-	-	-	-
Low-income subsidy	3	-	-	-	-	-
No insurance/insurance type	3	-	1	-	-	-
Indices/Scores						
Agency for Healthcare Research & Quality Socioeconomic Index	2	-	-	8	-	1
Area Deprivation Index	-	-	1	3	-	-
Diez-Roux score	-	-	1	-	-	-
Distressed Communities Index	-	-	2	-	-	-
Social Deprivation Index	1	-	-	-	-	-
Social Support						
Children	1	-	-	-	-	-
Home healthcare	1	-	-	-	-	-
Living alone	1	-	-	-	-	-
Marital status	4	-	-	-	-	-
No social/emotional support	-	-	1	-	-	-
Social Associations (# of membership organizations)	-	-	1	-	-	-
Functional Status-Related Factors						
Activities and Participation						
Activities of Daily Living[3]	3	-	-	-	-	-
Instrumental Activities of Daily Living[4]	2	-	-	-	-	-
Ambulatory function/impairment (e.g., gait speed, fall risk/scale)	3	-	-	2	-	-
Frailty	-	-	-	2	-	-
Body Function						
Cognitive function/impairment	1	-	-	2	-	-
Motor function	1	-	-	-	-	-
Functional Status						
3M™ Clinical Risk Group	1	-	-	-	-	-
Independent/Dependence	2	-	-	-	-	-

[3] Activities of Daily Living include eating, bathing, dressing, toileting, mobility, and grooming.

[4] Instrumental Activities of Daily Living are slightly more complex skills and include managing finances, handling transportation, shopping, preparing meals, using the telephone or other communication devices, managing medications, doing laundry, housework, and basic home maintenance.

Conclusion

NQF endorses performance measures that are intended for use in both performance improvement and accountability applications, such as public reporting and pay-for-performance. In this context, the overall performance of a healthcare entity is linked to payment and used to assist patients in making informed decisions about their care. These performance comparisons should be affected as little as possible by factors other than the quality of care, such as patient characteristics, including social and functional status-related factors.

Risk adjusting outcome performance measures to account for differences in patient characteristics is widely accepted, further ensuring fair, unbiased, and accurate measurement. However, for social and functional risk, there is variability in how this is done. Approaches to risk adjustment of social and functional risk ranges in the statistical models used and the steps taken to determine whether these factors are included in the overall risk model. Additionally, various data sources are used for social and functional risk factor adjustment, including community-level assessments of SES/SDS risk (e.g., AHRQ SES). However, challenges with data availability and accessibility exist for certain factors. For example, there is a paucity of functional status-related risk factors used within quality measure risk models compared to social risk factors. This may be due to the lack of an accepted approach to defining functional status and the limited availability of functional status data, which is largely sourced directly from the patient through survey instruments and/or assessments. For social risk factors at the patient-level, there is limited use of other sociodemographic data, such as education, employment, income, housing, etc. Again, this may largely be due to the lack of generalizable data that can be used within quality improvement mechanisms.

As the quality measurement enterprise continues to tie payment to quality of care provided, more guidance is needed for measure developers related to social and functional status-related risk adjustment. For the next phase of this work, NQF will develop TEP-informed technical guidance on the process of developing a risk adjustment model for outcome and resource use measures. However, limitations may still exist in the short-term due to potential limitations in data availability, as more generalizable data sources and definitions are needed for certain risk factors (e.g., functional status). Furthermore, measures interact with program design elements and there needs to be additional transparency within and across these programs due to the potential implications for risk adjustment.

References

1 Iezzoni LI, ed. *Risk Adjustment for Measuring Health Care Outcomes*. 4th ed. Chicago, IL: Health Administration Press; 2013.

2 Franks P, Fiscella K. Effect of Patient Socioeconomic Status on Physician Profiles for Prevention, Disease Management, and Diagnostic Testing Costs. *Medical Care*. 2002;40:717-724.

3 Blum AB, Egorova NN, Sosunov EA. Impact of Socioeconomic Status Measures on Hospital Profiling in New York City. *Circulation: Cardiovascular Quality and Outcomes*. 2014;7:391-397.

4 Bierman AS, Lawrence WF, Haffer SC, et al. Functional health outcomes as a measure of health care quality for Medicare beneficiaries. *Health Serv Res*. 2001;36(6 Pt 2):90-109.

5 Institute of Medicine (US) Committee on Quality of Health Care in America. *Crossing the Quality Chasm: A New Health System for the 21st Century*. Washington (DC): National Academies Press (US); 2001. http://www.ncbi.nlm.nih.gov/books/NBK222274/. Last accessed February 2021.

6 Analysis of Crosscutting Medicare Functional Status Quality Metrics Using the Continuity and Assessment Record and Evaluation (CARE) Item Set. ASPE. https://aspe.hhs.gov/basic-report/analysis-crosscutting-medicare-functional-status-quality-metrics-using-continuity-and-assessment-record-and-evaluation-care-item-set. Published June 13, 2015. Last accessed February 2021.

7 Bernheim SM, Parzynski CS, Horwitz L. Accounting for Patients' Socioeconomic Status Does not Change Hospital Readmission Rates. *Health Affairs*. 2016;35:1461-1470.

8 Chatterjee P, Werner RM. The Hospital Readmission Reduction Program and Social Risk. *Health Services Research*. 2019;54:324-326.

9 National Quality Forum. Evaluation of the NQF Trial Period for Risk Adjustment for Social Risk Factors. *July*. 2017. https://www.qualityforum.org/Publications/2017/07/Social_Risk_Trial_Final_Report.aspx.

10 Risk Adjustment for Socioeconomic Status or Other Sociodemographic Factors. *NQF*. 2014. https://www.qualityforum.org/WorkArea/linkit.aspx?LinkIdentifier=id&ItemID=77474.

11 NQF: Adjusting Measures for Social Risk Factors. http://www.qualityforum.org/Adjusting_Measures_for_Social_Risk_Factors.aspx#:~:text=NQF%3A%20Adjusting%20Measures%20for%20Social%20Risk%20Factors&text=The%20National%20Quality%20Forum's%20(NQF,a%20person's%20level%20of%20education. Last accessed January 2021.

12 International Classification of Functioning, Disability and Health (ICF). https://www.who.int/classifications/international-classification-of-functioning-disability-and-health.

13 *Classifying and Reporting Functional Status*. National Committee on Vital and Health Statistics (NCVHS), Subcommittee on Populations https://www.ncvhs.hhs.gov/wp-content/uploads/2017/08/010617rp.pdf.

14 Accounting for social risk factors in Medicare payment Identifying social risk factors. In: *National Academies of Sciences Engineering, and Medicine*. Washington, DC: The National Academies Press; 2016. https://www.ncbi.nlm.nih.gov/books/NBK338754/pdf/Bookshelf_NBK338754.pdf.

15 Introduction: What is MeSH? https://www.nlm.nih.gov/bsd/disted/meshtutorial/introduction/index.html.

16 A.S.P.E. *Social Risk Factors and Performance Under Medicare's Value-Based Payment Programs*. https://aspe.hhs.gov/system/files/pdf/253976/RTCAppendices.pdf.

17 A.S.P.E. *Social Risk Factors and Performance Under Medicare's Value-Based Payment Programs*. https://aspe.hhs.gov/system/files/pdf/263676/Second-IMPACT-SES-Report-to-Congress.pdf.

18 O'Malley AJ, Zaslavsky AM, Elliott MN, et al. Case-mix adjustment of the CAHPS Hospital Survey. *Health Serv Res*. 2005;40(6 Pt 2):2162-2181.

19 *Measures Management System. CMS*. https://www.google.com/url?sa=t&rct=j&q=&esrc=s&source=web&cd=&ved=2ahUKEwizi4vEnPjsAhVooXIEHYJoDqsQFjACegQIAxAC&url=https%3A%2F%2Fwww.cms.gov%2FMedicare%2FQuality-Initiatives-Patient-Assessment-Instruments%2FMMS%2FDownloads%2FRisk-Adjustment.pdf&usg=AOvVaw1J7CVsRimaXQA95_TA4QgQ.

20 Greenhalgh. Effectiveness and efficiency of search methods in systematic reviews of complex evidence: audit of primary sources. *BMJ*. 2005;331(7524):1064-1065.

21 *Measure Sets and Measurement Systems: Multistakeholder Guidance for Design and Evaluation*. National Quality Forum; 2020. https://www.qualityforum.org/WorkArea/linkit.aspx?LinkIdentifier=id&ItemID=93455.

22 *Office of the Assistant Secretary for Planning and Evaluation Report to Congress 2020, "Social Risk Factors and Performance in Medicare's Value-Based Purchasing Program*.

23 Pencina MJ, D'Agostino RB, Demler OV. Novel metrics for evaluating improvement in discrimination: net reclassification and integrated discrimination improvement for normal variables and nested models. *Stat Med*. 2012;31(2):101-113.

24 Pencina MJ, D'Agostino RB, Steyerberg EW. Extensions of net reclassification improvement calculations to measure usefulness of new biomarkers. *Stat Med*. 2011;30(1):11-21.

25 Pencina MJ, D'Agostino RB, D'Agostino RB, et al. Evaluating the added predictive ability of a new marker: from area under the ROC curve to reclassification and beyond. *Stat Med*. 2008;27(2):157-172; discussion 207-212.

26 Bonito A, Carpenter L, Eicheldinger C, et al. *Creation of New Race-Ethnicity Codes and Socioeconomic Status (SES) Indicators for Medicare Beneficiaries*. RTI International, Agency for Healthcare Research and Quality

27 The Spaces Between Us: The Evolution of American Communities in the New Century. *Economic Innovation Group*. https://eig.org/dci. Last accessed February 2021.

28 NQF #3474 Measure Submission Documents. http://www.qualityforum.org/ProjectTemplateDownload.aspx?SubmissionID=3474. Last accessed January 2021.

29 Medicare 2020 Part C & D Star Ratings Technical Notes. https://www.cms.gov/Medicare/Prescription-Drug-Coverage/PrescriptionDrugCovGenIn/Downloads/Star-Ratings-Technical-Notes-Oct-10-2019.pdf.

30 Hospital Readmissions Reduction Program (HRRP). https://www.cms.gov/Medicare/Medicare-Fee-for-Service-Payment/AcuteInpatientPPS/Readmissions-Reduction-Program.

31 Design for Nursing Home Compare Five-Star Quality Rating System: Technical Users' Guide. https://www.cms.gov/Medicare/Provider-Enrollment-and-Certification/CertificationandComplianc/downloads/usersguide.pdf.

32 *Home Health Quality Reporting Program Measure Calculations and Reporting User's Manual Version 1.0.*; 2019. https://www.cms.gov/Medicare/Quality-Initiatives-Patient-Assessment-Instruments/HomeHealthQualityInits/Downloads/Home-Health-QRP-QM-Users-Manual-V10-August-2019.pdf.

33 Medicare and Medicaid Programs; CY 2021 Home Health Prospective Payment System Rate Update, Home Health Quality Reporting Program Requirements, and Home Infusion Therapy Services and Supplier Enrollment Requirements; and Home Health Value-Based Purchasing Model Data Submission Requirements. Federal Register. https://www.federalregister.gov/documents/2020/11/04/2020-24146/medicare-and-medicaid-programs-cy-2021-home-health-prospective-payment-system-rate-update-home. Published November 4, 2020. Last accessed January 2021.

34 Claims Based ACH and ED Use Measures Technical Documentation and Risk Adjustment. https://www.cms.gov/Medicare/Quality-Initiatives-Patient-Assessment-Instruments/HomeHealthQualityInits/Downloads/Claims-Based-ACH-and-ED-Use-Measures-Technical-Documentation-and-Risk-Adjustment.zip.

35 *Evaluation of the Home Health Value-Based Purchasing (HHVBP) Model Third Annual Report*. Centers for Medicare & Medicaid Services; 2020. https://innovation.cms.gov/data-and-reports/2020/hhvbp-thirdann-rpt.

36 CMS ESRD Measures Manual for the 2017 Performance Period. https://www.cms.gov/Medicare/Quality-Initiatives-Patient-Assessment-Instruments/ESRDQIP/Downloads/ESRD-Manual-v20.pdf.

37 *CMS ESRD Measures Manual for the 2021 Performance Period.*; 2020. https://www.cms.gov/files/document/esrd-measures-manual-v60.pdf.

38 The Consumer Assessment of Healthcare Providers and Systems (CAHPS®) In-Center Hemodialysis Survey. https://ichcahps.org/. Last accessed February 2021.

39 *2019 Managed Long-Term Care Report*. New York State Department of Health https://www.health.ny.gov/health_care/managed_care/mltc/pdf/mltc_report_2019.pdf.

40 *A Path toward Value Based Payment: Annual Update New York State Roadmap for Medicaid Payment Reform*. New York State Department of Health; 2019. https://www.health.ny.gov/health_care/medicaid/redesign/dsrip/vbp_library/2020/docs/2019-09_final_vbp_roadmap.pdf.

41 Roux AVD, Merkin SS, Arnett D, et al. Neighborhood of Residence and Incidence of Coronary Heart Disease. *New England Journal of Medicine*. 2001;345(2):99-106.

42 Butler DC, Petterson S, Phillips RL, et al. Measures of Social Deprivation That Predict Health Care Access and Need within a Rational Area of Primary Care Service Delivery. *Health Services Research*. 2013;48(2pt1):539-559.

Appendices

Appendix A: TEP Members, Federal Liaisons, and NQF Staff

Committee Members

Philip Alberti, PhD *(Co-chair)*
Senior Director, Health Equity Research and Quality, Association of American Medical Colleges
Washington, District of Columbia

Karen Joynt Maddox, MD, MPH *(Co-chair)*
Assistant Professor, Washington University School of Medicine
St. Louis, Missouri

Arlene Ash, PhD
Professor and Division Chief, Biostatistics and Health Services Research, Department of Population and Quantitative Health Sciences, University of Massachusetts Medical School
Boston, Massachusetts

Patrick Campbell, PhD, PharmD, RPh
Senior Director, Measurement Outcomes Research, Pharmacy Quality Alliance
Alexandria, Virginia

Elizabeth Drye, MD, SM
Senior Director, Quality Measurement Programs**,** Yale-New Haven Hospital, Center for Outcomes Research and Evaluation (CORE)
New Haven, Connecticut

Marc Elliott, PhD, MA
Senior Principal Researcher, The RAND Corporation
Santa Monica, California

Rachel Harrington, PhD
Research Scientist, National Committee for Quality Assurance
Washington, District of Columbia

Bellinda King-Kallimanis, PhD, MSc.
Director of Patient-Focused Research, LUNGevity
Bethesda, Maryland

Vincent Liu MD, MS
Research Scientist, Kaiser Permanente Division of Research
Oakland, California

Danielle Lloyd, MPH
SVP Private Market Innovations & Quality Initiatives, America's Health Insurance Plans
Washington, District of Columbia

John Martin, PhD, MPH
Vice President, Data Science, Premier Healthcare Alliance

Charlotte, North Carolina

Shalini Prakash, MS
Data Scientist, Washington Healthcare Authority
Olympia, Washington

Sandra Richardson, MS
Research Scientist; Director, Bureau of Quality Measurement of Special Populations
Albany, New York

David Shahian, MD
Vice President, Center for Quality and Safety, Massachusetts General Hospital
Boston, Massachusetts

Cristie Upshaw Travis, MSHHA
Chief Executive Officer, Memphis Business Group on Health (MBGH)
Memphis, Tennessee

Janice Tufte
Hassanah Consulting
Seattle, Washington

Katherine Vickery, MD, MSc.
Clinician-Investigator, Hennepin Healthcare
Minneapolis, Minnesota

Federal Liaisons

Shafa Al-Showk, PhD
Center for Medicare/Centers for Medicare & Medicaid Services (CM/CMS)

Joel Andress, PhD
Center for Clinical Standards and Quality/Centers for Medicare & Medicaid Services (CCSQ/CMS)

Craig Caplan, MA
Federal Office of Rural Health Policy
Health Resources and Service Administration (HRSA)

Sophia Chan, PhD, MPH
Center for Medicare/Centers for Medicare & Medicaid Services (CM/CMS)

Maushaumi (Mia) DeSoto, PhD
Center for Quality Improvement and Patient Safety
Agency for Healthcare Research and Quality (AHRQ)

Andy Frankos-Rey, MA
Center for Medicaid & CHIP Services/ Centers for Medicare & Medicaid Services (CMCS/CMS)

Sarah Gaillot, PhD
Center for Medicare/Centers for Medicare & Medicaid Services (CM/CMS)

David Nyweide, PhD
Center for Medicare & Medicaid Innovation/Centers for Medicare & Medicaid Services (CMMI/CMS)

Jesse Roach, MD
Center for Clinical Standards and Quality/Centers for Medicare & Medicaid Services (CCSQ/CMS)

Lok Wong Samson, PhD
Office of Health Policy
Assistant Secretary for Planning and Evaluation (ASPE)

Rachael Zuckerman, PhD
Office of Health Policy
Office of the Assistant Secretary for Planning and Evaluation (ASPE)

NQF Staff

Sheri Winsper, RN, MSN, MSHA
Senior Vice President, Quality Measurement

Sai Ma, PhD
Managing Director, Senior Technical Expert

Matthew K. Pickering, PharmD
Senior Director

Taroon Amin, PhD, MPH
Consultant

Katie Berryman, MPAP, PMD
Senior Project Manager

Janaki Panchal, MSPH
Manager

Hannah Ingber, MPH
Senior Analyst

Juanita Rogers, MS, CHES
Analyst

Appendix B: Glossary

Decomposition analysis seeks to examine the impact of factors in the risk model. In the context of social/functional risk, the goal of this analysis is generally to disentangle the patient-level and hospital-level risk factors in the risk adjustment models. If both the patient-level and hospital-level effect are significant, that indicates that both are associated with increased risk of the outcome being measured.

Clinical adjustment refers to adjustment for only those physiological and psychiatric attributes, which at certain levels may be associated with an increased risk of certain diseases or death.[10]

Functional status is variously defined in the health field. Generally, functional status refers to an attribute that assesses how a health condition has impacted an individual's body function, body structures, and ability to participate in activities and complete basic daily tasks.[12] Functional status covers both the individual carrying out activities of daily living and the individual participating in life situations and society.[13] This includes 1) basic physical and cognitive activities such as walking or reaching, focusing attention, and communicating, as well as the routine activities of daily living, including eating, bathing, dressing, transferring, and toileting; and 2) life situations such as school or play for children and, for adults, work outside the home or maintaining a household. Furthermore, functional limitations occur when a person's capacity to carry out such activities or performance of such activities is compromised due to a health condition or injury and is not compensated by environmental factors (including physical, social, and attitudinal factors). Functional status encompasses the whole person and is affected by physical, developmental, behavioral, emotional, social, and environmental conditions.

Net reclassification index (NCI) is an index measure of how well a new model reclassifies subjects compared with an old model (correct versus incorrect changes in prediction for cases and controls separately). Therefore, it is a combination of these proportions with a maximum value of two.

Outcome will be used broadly to include the following types of outcomes relevant to performance measurement: quality outcomes of health outcome (e.g., mortality), intermediate clinical outcome (e.g., BP < 140/90), and economic outcomes of cost and resource use.

Risk adjustment refers to statistical methods to control or account for patient-, facility-, and/or community-level factors when computing performance measure scores; methods include modeling techniques, indirect standardization, or direct standardization. These methods can be used to produce a ratio of observed-to-expected, a risk-adjusted rate, or another estimate of performance.

Social risk factors are the social conditions associated with health (e.g., age, race, ethnicity, cultural context, gender, social relationships, residential and community context), including socioeconomic status (e.g., income, education, occupation).[14]

Social or functional status-related risk adjustment refers to statistical adjustment for sociodemographic and/or function status-related variables.

Appendix C: Data Source Descriptions

- **3M™ Clinical Risk Group**
  - The 3M Clinical Risk Groups (CRGs) are a population classification system that uses inpatient and ambulatory diagnosis and procedure codes, pharmaceutical data, and functional health status to assign each individual to a single, severity-adjusted group
  - Information on CRG access and use can be found at https://www.3m.com/3M/en_US/health-information-systems-us/drive-value-based-care/patient-classification-methodologies/crgs/
- **American College of Surgeons National Surgical Quality Improvement Program (ACS-NSQIP)**
  - The ACS-NSQIP is a validated, institution-based, multispecialty, surgical registry of patient risk factors and 30-day postoperative outcomes (https://www.facs.org/Quality-Programs/ACS-NSQIP).
- **American Community Survey**
  - The American Community Survey (ACS) is a nationwide survey designed to provide communities with reliable and timely social, economic, housing, and demographic data every year. The ACS Summary File is a set of comma-delimited text files that contain all of the Detailed Tables for the ACS data releases.
  - The ACS Summary File is located on the Census Bureau's file transfer protocol (FTP) server and is free to download (http://www.census.gov/acs).
- **Agency for Healthcare Research & Quality Socioeconomic Status (SES) Index**
  - The Agency for Healthcare Research and Quality (AHRQ) SES Index is derived from the American Community Survey (ACS) census block group level data and linked to a patient's ZIP code.[26]
- **Area Deprivation Index**
  - The Area Deprivation Index (ADI) is based on a measure created by the Health Resources & Services Administration (HRSA) over two decades ago for primarily county-level use, but refined, adapted, and validated to the Census block group/neighborhood level. It allows for rankings of neighborhoods by socioeconomic disadvantage in a region of interest (e.g., at the state or national level). It includes factors for the theoretical domains of income, education, employment, and housing quality. It can be used to inform health delivery and policy, especially for the most disadvantaged neighborhood groups.
  - The ADI project was supported by the National Institute on Aging of the National Institutes of Health, the National Institute On Minority Health And Health Disparities of the National Institutes of Health, and the University of Wisconsin School of Medicine and Public Health Department of Medicine.
  - Information on ADI access and use can be found at https://www.neighborhoodatlas.medicine.wisc.edu/
- **Area Health Resources File**
  - The Area Health Resource File (AHRF) compiles information from more than 50 databases and other sources to provide comprehensive county-level information on a variety of healthcare utilization, health professions and facilities, environmental, and socio-demographic topics. The basic file contains a variety of geographic descriptors and codes that enable aggregation of county-level data into geographic groupings and to link to other files.

  - The AHRF is maintained by the Health Resources and Services Administration (HRSA) and can be accessed from HRSA's website at https://data.hrsa.gov/data/download.
- **California Office of Statewide Health Planning and Development Data**
  - The California Office of Statewide Health Planning and Development (OSHPD) provides confidential patient-level data sets, consisting of patient discharge, ambulatory surgical center data, emergency department services, and birth and death data.
  - Information on access and use can be found at https://oshpd.ca.gov/data-and-reports/request-data/
- **Centers for Medicare & Medicaid Services Provider of Services (POS) Files**
  - The POS file contains data on characteristics of hospitals and other types of healthcare facilities, including the name and address of the facility and the type of Medicare services the facility provides, among other information. The data are collected through the Centers for Medicare & Medicaid Services (CMS) Regional Offices.
  - The POS is free and can be accessed at https://www.cms.gov/Research-Statistics-Data-and-Systems/Downloadable-Public-Use-Files/Provider-of-Services.
- **Chronic Condition Data Warehouse**
  - The Chronic Condition Data Warehouse (CCW) is a research databased with Medicare and Medicaid beneficiary, claims, and assessment data linked by beneficiary across the continuum of care.
  - CCW contains the following 100% Medicare files for years 1999 – 2018; 100% Medicare Encounter records for years 2015-2018; 100% Medicaid files for years 1999 – 2018; and 100% Part D Prescription Drug Event data for years 2006 - 2019:
  - CCW data files may be requested for any of the predefined chronic condition cohorts, or users may request a customized cohort(s) specific to research focus areas
  - Information on access and use can be found at https://www2.ccwdata.org/web/guest/request-data.
- **County Health Rankings & Roadmaps Program**
  - The County Health Rankings & Roadmaps (CHR&R) program is a collaboration between the Robert Wood Johnson Foundation, its sponsor, and the University of Wisconsin Population Health Institute.
  - The CHR&R program provides rankings to measure the health of nearly every county in all 50 states and report the performance of theses counties across a range of more than 30 health measures.
  - More information can be found at http://www.countyealthrankings.org/.
- **Diez-Roux Score**
  - The Diez-Roux neighborhood score is composed of six measures: the median household income; the median home value; the percentage of the population that completed high school; the percentage that completed college; the percentage with capital gains, dividend, or interest income; and the percentage with professional occupations. Each is normalized and summed. Higher scores correspond to higher SES neighborhoods[41].
- **Distressed Communities Index**
  - The Economic Innovation Group Distressed Communities Index (DCI) is available for all zip codes with more than 500 residents, which captures 99 percent of the American population. It is a composite score based on seven metrics: no high school degree, housing vacancy rate, adults not working, poverty rate, median income ratio, change in employment, and change in business establishments.

  - The 7 evenly weighted variables are used to calculate a zip code's rank compared with its geographic peers and then normalized to obtain a raw distress score that ranges from 0 (no distress) to 100 (severe distress). The seven SES indicators were obtained from the five-year estimates from the American Communities Survey and the Census Bureau County and ZIP Code Business Patterns.
  - Information on access and use can be found at https://eig.org/dci
- **Electronic Quality Improvement Platform for Plans and Pharmacies (EQuIPP®)**
  - The EQuIPP® database contains medication-related performance information for pharmacies covering 11.7 million Medicare beneficiaries
  - More information can be found at https://www.equipp.org/.
- **Emergency Department Benchmarking Alliance Database**
  - The Emergency Department Benchmarking Alliance (EDBA) provides its members with an independent database of demographic and performance metrics.
  - More information can be found at https://www.edbenchmarking.org/.
- **Health and Retirement Study-Medicare Linked Dataset**
  - The Health and Retirement Study (HRS)-Medicare linked dataset includes HRS survey information linked to CMS claims and assessment data for the HRS study population. The HRS has been fielded since 1992 and surveys more than 30,000 people age 50 and older. The HRS data include demographic and background information, including household; physical and mental health; cognition, functional limitations; employment; disability; health insurance; assets and income; and wills, including advanced directives.
  - Information on HRS data access and use can be found at https://www.resdac.org/cms-data/files/hrs-medicare.
- **Healthcare Cost and Utilization Project State Inpatient Databases**
  - The State Inpatient Databases (SID) are part of the family of databases and software tools developed for the Healthcare Cost and Utilization Project (HCUP) through a federal-state-industry partnership and sponsored by the Agency for Healthcare Research and Quality (AHRQ). The SID includes inpatient discharge records from community hospitals in that State.
  - The SID contain the universe of the inpatient discharge abstracts in participating states, translated into a uniform format to facilitate multi-state comparisons and analyses. The SID encompass about 97 percent of all U.S. community hospital discharges. Some states include discharges from specialty facilities, such as acute psychiatric hospitals.
  - The SID contain a core set of clinical and nonclinical information on all patients, including individuals covered by Medicare, Medicaid, or private insurance, as well as those who are uninsured.
  - Information on SID access and use can be found at https://www.hcup-us.ahrq.gov/sidoverview.jsp.
- **MarketScan© Commercial Claims and Encounter Data**
  - The MarketScan databases are a family of administrative claims databases that contain data on inpatient and outpatient claims, outpatient prescription claims, clinical utilization records, and healthcare expenditures.
  - Information on MarketScan access and use can be found at https://www.ibm.com/products/marketscan-research-databases/purchase.
- **Medicare Current Beneficiary Survey**

  - The Medicare Current Beneficiary Survey (MCBS) is a continuous, multipurpose survey of a nationally representative sample of the Medicare population, funded by the Office of Enterprise Data and Analytics (OEDA) of CMS.
  - The MCBS collects comprehensive data on beneficiaries' health insurance coverage, healthcare utilization and costs, access to care, and satisfaction with care, as well as special interest topics including drug coverage, knowledge about the Medicare program, and housing characteristics.
  - Information on MCBS access and use can be found at https://www.cms.gov/Research-Statistics-Data-and-Systems/Research/MCBS.
- **Medicare Fee-for-Service (FFS) Claims**
  - Medicare FFS claims (Inpatient, Outpatient, Skilled Nursing Facility, Hospice, Home Health, Carrier, Durable Medical Equipment), as well as the enrollment data.
  - Information on Medicare FFS data access and use can be found at https://www.resdac.org/research-identifiable-files-rif-requests
- **Medicare's Master Beneficiary Summary File (MBSF)**
  - The MBSF base segment includes beneficiary enrollment information, (Parts A, B, C, and D). Medicare Advantage (Part C) and the Prescription Drug Program (Part D) plan enrollment information is included.
- **Minimum Data Set (MDS)**
  - The MDS standardized, primary screening and assessment tool of health status which forms the foundation of the comprehensive assessment for all residents of long-term care facilities certified to participate in Medicare or Medicaid. The MDS contains items that measure physical, psychological, and psycho-social functioning.
  - Information on the MDS can be found at https://www.cms.gov/Medicare/Quality-Initiatives-Patient-Assessment-Instruments/NursingHomeQualityInits/NHQIQualityMeasures.
- **Outcome and Assessment Information Set**
  - The Outcome and Assessment Information Set (OASIS) is a group of standard data elements developed, tested, and refined over two decades through a research and demonstration program funded primarily by the Centers for Medicare & Medicaid Services, with additional funding from the Robert Wood Johnson Foundation and the New York State Department of Health. OASIS data elements were designed to enable systematic comparative measurement of home healthcare patient outcomes at two points in time.
  - The most recent dataset is OASIS-D. Information on OASIS can be found at: https://www.cms.gov/Medicare/Quality-Initiatives-Patient-Assessment-Instruments/HomeHealthQualityInits/HHQIOASISUserManual
- **Social Deprivation Index**
  - The Social Deprivation Index (SDI) summarizes seven sociodemographic measures taken from the US Census American Community Survey. The SDI was developed through a factor analysis of the percentage of the population that lives in poverty, percentage with <12 years of education, percentage of single-parent households, percentage living in rented housing units, percentage living in overcrowded housing units, percentage of households without a car, and percentage of non-employed adults under 65 years of age[42].
- **Society of Thoracis Surgeons/American College of Cardiology (STS/ACC) TVT Registry**

  - The STS/ACC TVT Registry™, created by a collaboration between the Society for Thoracic Surgeons (STS) and the American College of Cardiology (ACC), monitors patient safety and real-world outcomes related to transcatheter valve replacement and repair procedures—emerging treatments for valve disease patients.
  - Information on STS/ACC TVT Registry™ access and use can be found at https://www.ncdr.com/WebNCDR/tvt/publicpage/research
- **United States Census**
  - The Census Bureau collects data about the economy and the people living in the United States from many different sources. Some data are collected from respondents directly (including businesses) through the censuses and through surveys. Data are also collected from federal, state, and local governments, as well as some commercial entities. Information on Census data access and use can be found at: https://www.census.gov/about/adrm/linkage/guidance.html
- **Vascular Quality Initiative (VQI) Data Registry**
  - The VQI is a national cooperative quality improvement initiative developed to prospectively collect data and outcomes for patients undergoing vascular surgical procedures. It consists of >1,300 physicians from >300 academic and community medical centers across the United States.
  - Data are physician-reported at the time of operation and include preoperative, intraoperative, and in-hospital postoperative details. Follow-up data are entered at ~one year postoperatively. All information is sent to a central data repository where it is aggregated and audited. Research analysts are blinded to patient, surgeon, and hospital identity.
  - Information on VQI Data Registry access and use can be found at https://www.vqi.org/data-analysis/

Appendix D: Literature Review Summary Table

Please refer to Sheet 1 of Excel file, which can be found here (<insert link to NQF site>)

Appendix E: Measure Review Summary Table

Please refer to Sheet 2 of Excel file, which can be found here (<insert link to NQF site>)

Appendix F: Program Review Summary Table

Please refer to Sheet 3 of Excel file, which can be found here (<insert link to NQF site>)