

REPORT

FINAL REPORT

Proposed Methods for Developing and Testing Risk- and Reliability- Adjustment Models for HCBS Composite Measures

February 12, 2015

Alex Bohl
Mariel Finucane
Jessica Ross
Sheng Wang
Dejene Ayele

Submitted to:

U.S. Department of Health and Human Services
Centers for Medicare & Medicaid Services CMS/OAGM/AGG/DCRG
Mailstop: C2-21-15 7500 Security Blvd.
Baltimore, MD 21244-1850
Project Officer: Effie George
Contract Number: HHSM-500-2010-000261/HHSM-500-T0010

Submitted by:

Mathematica Policy Research
955 Massachusetts Avenue
Suite 801
Cambridge, MA 02139
Telephone: (617) 491-7900
Facsimile: (617) 491-8044
Project Director: Carol Irvin
Reference Number: 40137.c0f

This page left blank for double-sided copying.

ACKNOWLEDGEMENTS

A number of people at Mathematica and other organizations made important contributions to this report. We would like to thank our government task lead, D.E.B. Potter, of the Office of Disability, Aging, and Long-Term Care Policy in the Office of the Assistant Secretary for Planning and Evaluation. We would also like to thank our project officer, Effie George, and our deputy project officer, Michael Smith, both from the Division of Community Systems Transformation, Disabled & Elderly Health Programs Group, Center for Medicaid & CHIP Services at the Centers for Medicare & Medicaid Services (CMS). In addition, we would like to thank the members of the Agency for Healthcare Research and Quality (AHRQ) Quality Indicator team for their technical contributions. Finally, we wish to acknowledge Eric Schone for his significant contributions to this report, which benefited greatly from his statistical expertise and overall guidance.

This page left blank for double-sided copying.

CONTENTS

I	INTRODUCTION.....	1
II	DATA, MEASURE DEFINITIONS, AND MODELING FRAMEWORK.....	5
	A. Data and analytic populations	5
	B. Measure definitions	7
	1. Denominator (analytic population)	7
	2. Numerator	7
	3. Observed (unadjusted) rates	7
	C. Analytic challenges	8
	D. Role of the TEP	9
III	TECHNICAL APPROACH TO RISK ADJUSTMENT.....	11
	A. Candidate risk-adjustment models	11
	B. Selection of person-level risk factors.....	11
	C. Consideration of “HCBS exposure”	12
	D. State effects.....	13
	E. Risk-adjustment model development process and diagnostics	13
IV	TECHNICAL APPROACH TO RELIABILITY ADJUSTMENT.....	15
	A. Candidate approaches to reliability adjustment.....	15
	B. Specifying the distribution under which reliability adjustment is induced.....	15
	C. Reliability-adjustment model validation	16
V	COMPARISON OF STATE-LEVEL RESULTS.....	17
	A. Standardization.....	17
	B. Common methods for evaluating performance	17
	C. Setting an appropriate benchmark	19
VI	APPLY THE MEASURES TO RELEVANT SUBPOPULATIONS.....	21
VII	NEXT STEPS.....	23
	REFERENCES.....	25
	APPENDIX A: METHODS	

This page left blank for double-sided copying.

TABLES

I.1	Final AHRQ recommended measures	1
II.1	CCW candidate risk factors	5
V.1	Summary of common methods for evaluating performance	17
VII.1	Key dates and deliverables	24

FIGURES

II.1	Flowchart for HCBS composite development activities requiring TEP input	9
V.1	Simple ranking of state rates	18
V.2	State performance relative to a benchmark	18
V.3	Example of exceedance probability	19
A.1	Defining the Measure Denominator	A.4

This page left blank for double-sided copying.

I. INTRODUCTION

Section 6086(b) of the Deficit Reduction Act of 2005 directed the Agency for Healthcare Research and Quality (AHRQ) to develop “program performance indicators, client function indicators, and measures of client satisfaction” for Medicaid beneficiaries receiving home- and community-based services (HCBS) (U. S. Congress 2006). This directive led to the identification of 21 measure constructs, including avoidable hospitalizations and serious reportable events, which captured important dimensions of the quality of care delivered to HCBS recipients (AHRQ 2010). Subsequent work by AHRQ assessed whether 30 candidate indicators—adapted from the AHRQ Prevention Quality Indicators (PQIs) and Patient Safety Indicators (PSIs)—had the potential to adequately measure avoidable hospitalizations and serious reportable events among Medicaid beneficiaries receiving HCBS (Schultz et al. 2012). The final measure set included three composites and two individual measures:

- Composite: Ambulatory Care Sensitive Condition (ACSC) Chronic Conditions
- Composite: ACSC Acute Conditions
- Composite: ACSC Overall Conditions
- Individual Measure: Pressure Ulcers
- Individual Measure: Injurious Falls

The composite measures, which are adaptations of the AHRQ PQIs, report the rate of potentially avoidable hospitalization as a result of either chronic or acute conditions, as shown in Table I.1. These PQI measures monitor the occurrence of events resulting in hospitalization that should rarely occur when high quality outpatient care is provided, and as such, have been recognized as highly relevant to the HCBS community by several expert panels (Schultz et al. 2012; Davies et al. 2009).

Table I.1. Final AHRQ recommended measures

HCBS Composites	Component Indicators
ACSC Chronic Conditions Composite (PQI 92)	<ol style="list-style-type: none"> 1. Diabetes, short-term complications (PQI 1) 2. Diabetes, long-term complications (PQI 3) 3. COPD (PQI 5) 4. Hypertension (PQI 7) 5. CHF (PQI 8) 6. Angina without procedure (PQI 13) 7. Uncontrolled diabetes (PQI 14) 8. Adult asthma (PQI 15) 9. Lower extremity amputations among people with diabetes (PQI 16)
ACSC Acute Conditions Composite (PQI 91)	<ol style="list-style-type: none"> 1. Dehydration (PQI 10) 2. Bacterial Pneumonia (PQI 11) 3. Urinary Tract Infection (PQI 12)
ACSC Overall Composite (PQI 90)	All components from both the ACSC Chronic Conditions and ACSC Acute Conditions composites

Table I.1 (continued)

Source: Adapted from Schultz, E., S. Davies, and K. McDonald. "Development of Quality Indicators for Home and Community-Based Services Population: Technical Report." June 2012.

ACSC = Ambulatory Care Sensitive Condition; AHRQ = Agency for Healthcare Research and Quality; CHF = Congestive Heart Failure; COPD = Chronic Obstructive Pulmonary Disease; PSI = Patient Safety Indicator; PQI = Prevention Quality Indicator.

The three ACSC composites have the potential to provide important information about ACSC events and ACSC rates among Medicaid beneficiaries using HCBS at the national, state, program, or plan level. However, the original specifications developed by AHRQ did not include a robust methodology for accounting for case-mix differences or statistical uncertainty, which is necessary for comparing the quality of care delivered to HCBS users across states or other entities. To address this gap, the Centers for Medicare & Medicaid Services (CMS), AHRQ, and the Office of the Assistant Secretary for Planning and Evaluation (ASPE) directed Mathematica Policy Research to develop risk-adjustment methodology for these measures.

Mathematica developed models to risk-adjust these composites in 2013 (Ross and Bohl 2013). At the conclusion of this work, five areas for future refinement were identified:

- Address risk-adjustment model fit issues (for example, overprediction of events)
- Establish methods to make inferences about state HCBS composite values relative to benchmarks
- Incorporate reliability adjustment to improve the statistical precision of estimates for small states or small populations
- Explore strategies to produce stratified rates for meaningful subpopulations
- Incorporate newly available populations (for example, Medicaid-only HCBS recipients) and risk factors (for example, disability indicators) into the model

Since that time, the data needed to assess key risk factors have also become available, the science of modeling has progressed, and the focus of this work has shifted toward using the composites to make comparisons between state Medicaid programs. As a result, CMS has directed Mathematica to refine the risk-adjustment models developed previously and determine appropriate methods for facilitating state-level comparisons.

The purpose of this report is to describe the methods and processes to develop and test the risk-adjusted HCBS composites, with the ultimate goal of the finalized HCBS composites being to compare state-level rates for selected HCBS populations to relevant benchmarks. The HCBS composites are intended to assess the quality of care for HCBS recipients under a shared accountability framework: the measures profile the experience of the HCBS population and will reflect the care delivered by all providers (not just HCBS providers). These comparisons are intended to guide HCBS quality improvement efforts led by HCBS stakeholders such as CMS and state Medicaid offices. With this goal in mind, the remainder of the report addresses:

- Data sources, populations, and measure definitions
- Analytic challenges
- Technical approach to risk adjustment
- Technical approach to reliability adjustment
- Comparisons of state-level performance
- Subpopulation of interest: Money Follows the Person (MFP) participants
- Next steps

This report will focus on the three composite measures, but the methods described could also be applied to the pressure ulcer measure.

This page left blank for double-sided copying.

II. DATA, MEASURE DEFINITIONS, AND MODELING FRAMEWORK

A. Data and analytic populations

The development populations for this work will include Medicaid beneficiaries using HCBS in 2009 or 2010. This population includes HCBS recipients who are only enrolled in Medicaid, as well as those eligible for both Medicare and Medicaid (Medicare-Medicaid eligible or MME). The data sources and methods used to define these cohorts are identical to those used for the preliminary phase of this work (described in Appendix A), with the following exceptions:

- The 2009 HCBS user population now includes data from 50 states, whereas previously data from only 43 states were available
- The 2010 HCBS user population now includes data from 49 states, whereas previously data from only 30 states were available

In addition to including data from nearly all states, the current work will also consider newly available Chronic Condition Warehouse (CCW) information on 15 disability-related conditions, 9 mental illnesses, and 2 substance abuse conditions as candidate risk factors, whereas previously information on only 27 chronic conditions was available (Table II.1). The newly available CCW risk factors are binary flags, but we will ask the technical expert panel (TEP) whether it is appropriate to create composite risk factor measures from these conditions.

Table II.1. CCW candidate risk factors

Chronic Conditions
1. Alzheimer's Disease
2. Alzheimer's Disease and Related Disorders or Senile Dementia
3. Acute Myocardial Infarction
4. Anemia
5. Asthma
6. Atrial Fibrillation
7. Breast Cancer
8. Colorectal Cancer
9. Endometrial Cancer
10. Lung Cancer
11. Prostate Cancer
12. Cataract
13. Heart Failure
14. Chronic Kidney Disease
15. Chronic Obstructive Pulmonary Disease and Bronchiectasis
16. Depression
17. Diabetes
18. Glaucoma
19. Hip/Pelvis Fracture

Chronic Conditions

20. Hyperlipidemia
21. Hypertension
22. Benign Prostatic Hyperplasia
23. Acquired Hypothyroidism
24. Ischemic Heart Disease
25. Osteoporosis
26. Rheumatoid Arthritis/Osteoarthritis
27. Stroke/Transient Ischemic Attack

NEW: Disability-Related Conditions

1. Autism Spectrum Disorders
2. Cerebral Palsy
3. Cystic Fibrosis and Other Metabolic Developmental Disorders
4. Epilepsy
5. Intellectual Disabilities and Related Conditions
6. Learning Disabilities
7. Mobility Impairments
8. Multiple Sclerosis and Transverse Myelitis
9. Muscular Dystrophy
10. Other Developmental Delays
11. Sensory: Deafness and Hearing Impairment
12. Sensory: Blindness and Visual Impairment
13. Spina Bifida and Other Congenital Abnormalities of the Nervous System
14. Spinal Cord Injury
15. Traumatic Brain Injury and Nonpsychotic Mental Disorders due to Brain Damage

NEW: Mental Illnesses

1. Anxiety Disorders
2. Bipolar Disorder
3. Conduct Disorders and Hyperkinetic Syndrome
4. Depressive Disorders
5. Personality Disorders
6. Post-Traumatic Stress Disorders
7. Schizophrenia
8. Schizophrenia and Other Psychotic Disorders
9. Tobacco Use

NEW: Substance-Abuse Conditions

1. Alcohol Use
2. Substance Abuse

Source: Chronic Conditions Warehouse (<https://www.ccwdata.org/>).

The Medicare and Medicaid inpatient and administrative data available for analysis contain the information needed to specify the measure denominators and numerators, calculate observed (unadjusted) rates, and develop appropriate risk- and reliability-adjustment models, as described below.

B. Measure definitions

1. Denominator (analytic population)

The denominator for all three composite measures includes the total number of Medicaid beneficiaries either enrolled in or using HCBS 1915(c) waivers or those using state plan HCBS for at least one month during the period of interest. For these analyses, we will construct two analytic populations: one for Medicaid beneficiaries using HCBS in 2009, and the other for beneficiaries using HCBS in 2010. The 2009 and 2010 time periods represent the most recent years for which almost all states have available data, with data from 50 states available for 2009 and from 49 states available for 2010.

Several important exclusions are imposed on these populations, in accordance with the specifications developed by AHRQ. We excluded both Medicaid managed care and Medicare Advantage enrollees, because their claims either are not available or are not comparable to those for beneficiaries enrolled in fee-for-service. The population also excludes children by limiting the data to HCBS users who are age 18 or older. Finally, people with a record of HCBS enrollment only (that is, no observed HCBS claims) and at least one month with an institutional claim for long-term care were excluded. This step removes individuals who are enrolled in HCBS 1915(c) waivers but are only receiving institutional long-term services and supports during the period of interest.

Appendix A provides additional details on how we identified the HCBS recipient analytic population.

2. Numerator

For each composite measure, the numerator includes the total count of inpatient acute care hospital admissions with diagnosis or procedure codes meeting the criteria for any of the component measures (Table I.1). If an HCBS user experiences multiple qualifying hospital admissions during the period of interest, all of these admissions are counted in the numerator. This means that, although an HCBS user can only be counted once in the denominator, a user with more than one distinct hospital admission can contribute multiple times to the numerator.

In order to ensure that a hospital admission is counted only once, we excluded both transfers and duplicate records as described in Appendix A.

3. Observed (unadjusted) rates

The observed (unadjusted) composite rate for the time period of interest is calculated as the number of qualifying inpatient admissions divided by the number of HCBS users, e.g.,

$$\frac{\text{Number of qualifying inpatient admissions}}{\text{Total number of HCBS Users}} = \text{Rate of all events among all HCBS users}$$

This rate will include qualifying inpatient admissions from HCBS users who are admitted to the hospital once, as well as admissions from those who are admitted to the hospital multiple times during the period of interest. The rate defined in this way will be the primary focus of this risk- and reliability-adjustment work.

C. Analytic challenges

For each of these possible definitions, the unit of analysis is the person-year, meaning that the measure assesses the likelihood that the HCBS user will experience at least one qualifying inpatient admission during the calendar year (either 2009 or 2010). These units are then aggregated to create state-level results, which measure the likelihood that HCBS users *in a given state* will experience at least one qualifying inpatient admission during the calendar year.

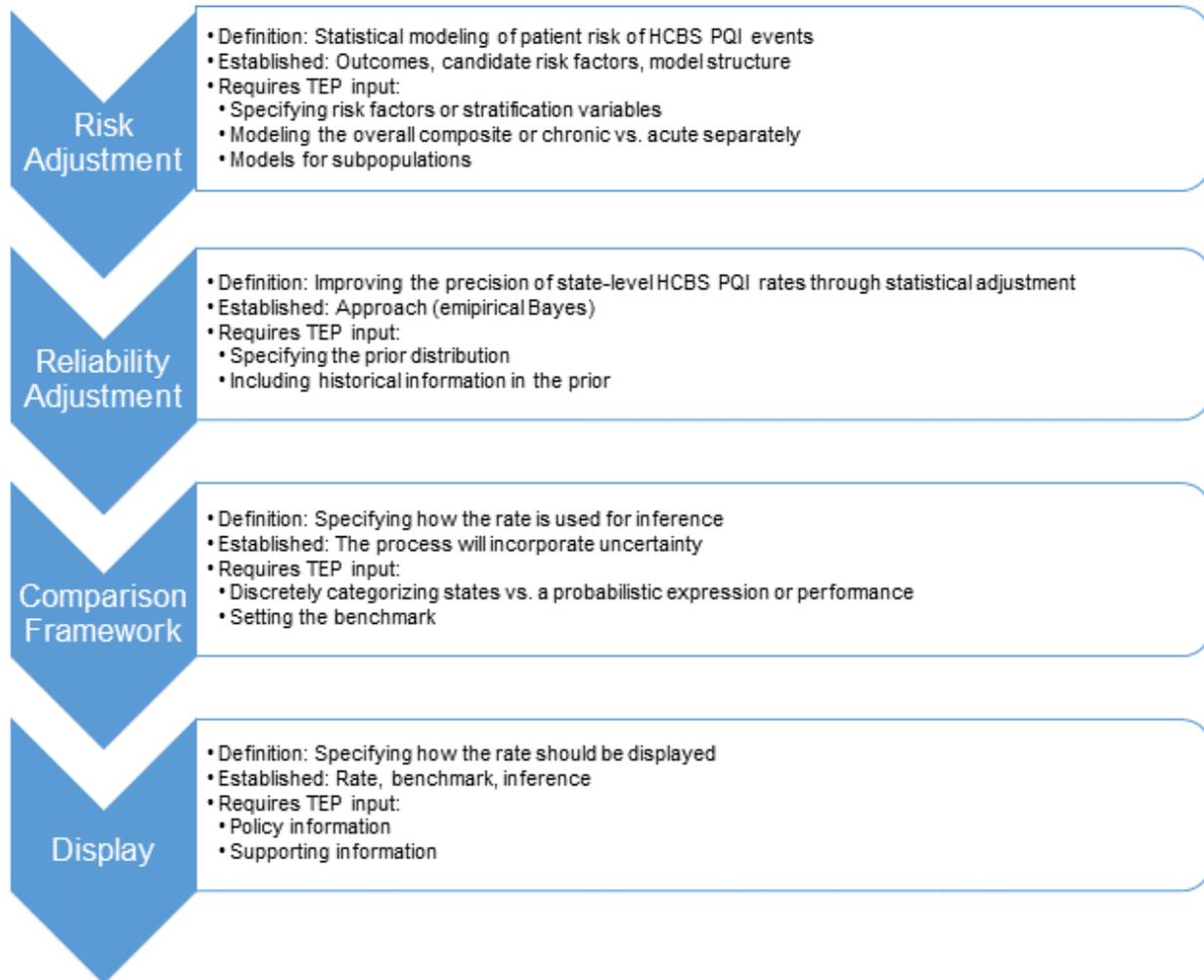
Although differences in observed state-level rates provide useful information, before assessing the quality of care delivered to HCBS recipients, it is important to account for differences in the following factors:

- **Person-level risk factors**, such as age or comorbidity status. These factors are defined at the beginning of the calendar year of analysis. This helps ensure that conditions that develop during the period of analysis and may be affected by quality of care are not considered—in other words, this definition ensures that risk factors are exogenous.
- **Exposure time**. If the outcome is influenced by the quality of care delivered to HCBS recipients, then the amount of exposure, or in this case HCBS use, also should be addressed. In this analysis, monthly indicators of HCBS use are available, as well as monthly enrollment in specific HCBS 1915(c) waivers.
- **Data quantity differences**. Even after accounting for all the above factors, the reliability of measure results is highly influenced by sample sizes. For example, the 2009 HCBS user population in California includes approximately 400,000 beneficiaries, while Tennessee has approximately 763 HCBS recipients. Therefore, Tennessee's rates are subject to random fluctuations to a greater extent than California's. For this reason, it's critical to account for large differences in sample sizes either by applying minimum case criteria or imposing reliability adjustment. Because the objective for this work is to create rates for all states, we focus on reliability adjustment rather than minimum case size thresholds.
- **State Medicaid eligibility and HCBS availability differences**. Individual states determine eligibility criteria for Medicaid. Each state also determines the HCBS available to Medicaid recipients through state plan or 1915 (c) waiver programs. On the one hand, Medicaid eligibility and HCBS availability are key determinants of a state's HCBS user population, and these differences might need to be accounted for through risk-adjustment. On the other hand, decisions on Medicaid eligibility and HCBS availability are *modifiable* and are therefore directly related to the quality of care for HCBS recipients; it might therefore be important not to obscure these differences. States may determine to change eligibility or HCBS availability to improve the quality of care for their HCBS population. We will bring this matter to the TEP for guidance.

D. Role of the TEP

Although the initial risk-adjustment work made significant progress on developing risk-adjustment models for the HCBS composite measures, input from a TEP is crucial for ensuring the utility of the final HCBS composites. A TEP consisting of HCBS providers and stakeholders, as well as researchers in the quality measurement field, will guide the HCBS composite development process. Figure II.1 describes the key areas for TEP input. The HCBS composite process involves four key activities, to be performed in this order: (1) risk adjustment, (2) reliability adjustment, (3) comparison framework, and (4) display. Within each of these activities, the TEP will provide input on specific analytic decisions. For example, while the outcome and distribution of the risk-adjustment models are specified, we require input on how to select and consider risk factors. In general, the downstream activities require more input than the upstream activities. This is because a group of experts provided previous feedback on the risk-adjustment process but was not asked to consider the other topics.

Figure II.1. Flowchart for HCBS composite development activities requiring TEP input



The remainder of this report details the technical approaches we will explore to address each of these issues and how we will evaluate their relative success.

This page left blank for double-sided copying.

III. TECHNICAL APPROACH TO RISK ADJUSTMENT

A. Candidate risk-adjustment models

The HCBS composite numerator represents a count—the number of hospitalizations meeting the measure inclusion criteria over a given calendar year. For a predictive model of this count, we will consider the following statistical distributions that can be used to model count data:

1. Poisson
2. Negative binomial
3. Zero-inflated Poisson
4. Zero-inflated negative binomial

Poisson and negative binomial regression models can both be used to fit count data. However, Poisson and negative binomial regression models differ in their assumptions about the conditional mean and variance of the counts. Poisson models assume that the conditional mean and variance of the distribution are equal. Negative binomial regression models do not assume an equal mean and variance, thereby accounting for over-dispersion in the data, which is when the variance is greater than the conditional mean (Osgood 2000; Paternoster and Brame 1997).

Choosing between Poisson and negative binomial models depends on the nature of the distribution of the counts—in this case, the counts of qualifying inpatient hospital admissions. Analysts commonly select negative binomial regression because the restrictive variance-equals-mean assumption of Poisson models often does not hold in health data. However, Poisson distributions have advantages under some circumstances. For example, when analyzing small samples, the negative binomial model is not recommended, so Poisson is the best available choice (Long 1997).

Zero-inflated regression is used to model count data with an excess of zero counts. Specifically, zero-inflated regression incorporates the assumption that the zeroes are generated by a separate process from the count values and that the zeroes can be modeled independently. The preliminary risk-adjustment work modeled the count of events using a negative binomial regression. In this phase of measure development, we will test zero-inflated versions of the Poisson and negative binomial models because the number of persons with zero HCBS composite events is greater than that predicted by standard count models.

B. Selection of person-level risk factors

During the preliminary phase of this work, Mathematica convened two conference calls to solicit feedback from experts on potential risk factors and our technical approach to risk adjustment. These calls included participants with expertise on the HCBS user population, risk-adjustment, and measure development in general. A number of attendees had participated in

AHRQ's original work to develop the HCBS measures.¹ The main points from these discussions were as follows:

- Risk adjustment by age, sex, and comorbidities has the advantage of simplicity and exogeneity.
- Potential risk factors should include characteristics that are predictive of hospital admissions in general, when possible.
- Capturing limitations in the capacity to perform activities of daily living (ADL) is important for risk adjustment, but it might not be possible to measure it for all HCBS users.
- Mental health conditions should be considered in the risk-adjustment model, either individually or as a composite.
- Medicaid eligibility criteria and waiver use or enrollment should not be used as a proxy for ADL/disability level due to heterogeneity in the definition of waiver programs across states.

This feedback was used to select the person-level risk factors that were included in the preliminary risk-adjustment models. However, when these models were developed, data for only 27 chronic conditions were available, whereas information for 15 disabilities, 9 mental illnesses, and 2 substance abuse conditions now can be considered (Table II.1). Although the above recommendations provide a foundation for our current work, the addition of 26 new candidate risk factors represents a substantial amount of new information. As a result, we will convene another expert panel to review the new potential risk factors and help determine which ones should be included in the model due to clinical significance.

Our proposed approach is to keep risk factors that are either statistically or clinically significant in the model, with the following exceptions (Gelman and Hill 2007):

1. A (clinically significant) risk factor will be removed if its coefficient is not statistically significantly different from zero and is judged to be multicollinear with other risk factors. Although this may be a concern to stakeholders interested in a certain risk factor, it is important to remove such nonsignificant risk factors, because their effects cannot be reliably estimated from the data.
2. A risk factor will be removed if it has a coefficient with a relatively large magnitude that is in the opposite direction of what is hypothesized.

C. Consideration of “HCBS exposure”

As discussed previously, it is also important to account for the different amount of “HCBS exposure”—the number of months that an individual is using HCBS or enrolled in an HCBS plan—in the model. In the preliminary risk-adjustment work, we modeled the effect of exposure by treating it as an “offset.” This specification enters exposure time in the model but constrains the coefficient on that variable to be one. However, this specification requires the assumption

¹ Comments from the expert panel for HCBS measure development are available at http://qualityindicators.ahrq.gov/downloads/resources/publications/2012/appendix_1b_details_of_expert_panel_calls.pdf

that there is a proportional relationship between the offset variable (months of HCBS) and the outcome (rate of inpatient admissions). If results from the updated analytic files do not meet this assumption, other options for addressing differences in HCBS exposure will be explored. These options may include adding months of HCBS as an explanatory variable, or imposing restrictions to exclude HCBS users with only small amounts of observation time (3 months or less) from the denominator. We will test these different treatments of the exposure variable in comparison with its original specification as an “offset.”

D. State effects

The focus of this work is to identify state-level differences or “effects” in the HCBS composite measure rates, after accounting for differences in person-level risk factors, exposure time, and other influences that may affect rates but are not directly related to the quality of care. To accomplish this goal, these state effects may be included in the model as fixed or random effects, or may be omitted from the model entirely. The differences between these options can be described as follows:

- **Fixed effect.** In this model, a fixed coefficient is estimated for each state, similar to how other independent variables’ coefficients are estimated. Although this approach yields unbiased estimates of the state-level effects of interest, these estimates may be highly imprecise and are therefore commonly stabilized using a subsequent reliability adjustment step.
- **Random effect.** In this model, a random intercept is estimated for each state. This approach yields stable estimates of state-level effects, and subsequent reliability adjustment is therefore not required. The stability of these random effect estimates comes at the cost of potential bias. In particular, the estimates will be biased if the state-level effects are correlated with other risk factors in the model. This bias can be mitigated by including state-level covariates that may help to explain this problematic correlation (Ash et al. 2012).
- **No effect.** If we believe that there are no state-level characteristics that affect the outcome but have not been included in the model, the modeling of a state effect may not be needed. However, a model that excludes state-level effects may not achieve full case-mix adjustment, due to possible confounding of the relationship of patient risk with the outcome (Ash et al. 2012).

Although the preliminary risk-adjustment models did not include state effects, we will explore all three options during this phase of the work.

E. Risk-adjustment model development process and diagnostics

As described above, for each of the overall, acute, and chronic HCBS composite measures, we will choose (1) a statistical distribution with which to model the count data, (2) risk factors to adjust for, (3) a method for accounting for HCBS exposure time, and (4) whether and how to include state effects. Throughout, our model development work will prioritize models with the greatest predictive validity (Ash et al. 2012), focusing on the model-building processes outlined in the hierarchical modeling literature (Gelman and Hill 2007).

The model development process for each composite measure will begin with an evaluation of the overdispersion parameter estimated in the Poisson/zero-inflated Poisson regression model. An overdispersion parameter value much larger than one would be a strong indicator of an overdispersion issue in the model, which implies greater variability than would be expected and a potentially poor model fit. Negative binomial models are less susceptible to this issue because the presence of a second free parameter can account for additional variance. In addition, we will use the Vuong Likelihood Ratio Test to examine whether the zero-inflated model is a significant improvement over the regular (non-inflated) model.

Once the appropriate statistical distribution has been determined, person-level risk factors will be chosen as described in Section B, above. This model selection process will also entail choosing a method with which to account for HCBS exposure time. As described in Section C, above, if model diagnostics suggest that exposure time cannot simply be treated as an offset, then other options for addressing differences in HCBS exposure will be explored.

Analyses used to aid in these choices among distributions, risk factors, and exposure adjustments will include:

1. Comparison of Akaike information criterion (AIC) and Bayesian information criterion (BIC). The AIC and BIC consider the number of parameters, sample size, and maximum likelihood to assess model quality, with lower results indicating better model fit.
2. Consideration of scaled deviance and Pearson's chi-square statistic in models that either include state fixed effects or have no state effect. The scaled version of both of these statistics, under certain regularity conditions, has a limiting chi-square distribution, with degrees of freedom equal to the number of observations minus the number of parameters estimated. The scaled version can be used as an approximate guide to the goodness of fit of a given model.

After the optimal base model has been developed for each composite measure, we will determine whether and how to include state effects in the model. To assess the complex bias/variance tradeoffs entailed in this decision, we will use the simulation approach described by Clark and Linzer (2014), with the goal of minimizing the model's mean square error, a comprehensive measure of model accuracy. If a random effects approach is indicated, we will (1) consider including state-level covariates to mitigate bias, and (2) use posterior predictive checking as described below in Section IV.C to ensure that the model is providing a good fit to the data.

IV. TECHNICAL APPROACH TO RELIABILITY ADJUSTMENT

A. Candidate approaches to reliability adjustment

Reliability adjustment is the process of removing statistical “noise” or random error from measure results in order to produce more accurate comparisons between entities of interest such as states. The general approach is to shrink risk-adjusted rates toward an overall mean, with the degree of shrinkage depending on the amount of variability in the data. In particular, imprecise risk-adjusted rates (that is, those with large standard errors, often due to small sample size) will be shrunken toward the overall mean to a greater degree. The gain in precision achieved by reliability adjustment can be substantial, yielding estimates that are more stable over time and that suffer less from regression to the mean.

We will consider the following candidate approaches to estimating reliability-adjusted HCBS composite measures for each state:

1. **A unified model.** Including state-level random effects in the risk-adjustment model induces reliability adjustment, and a separate reliability-adjustment model is therefore not necessary. Unified models are common for hospital mortality and readmission models (Ash et al. 2012). Important statistical developments have occurred through hospital quality measure efforts, but these condition-specific mortality and readmission measures apply a unified logit model, whereas the HCBS composites require a slightly more complicated count model. The initial phase of developing HCBS composite risk-adjustment models found that the unified model has computational challenges when analyzing count data.
2. **A two-stage approach.** In a risk-adjustment model that does not include state-level random effects, reliability adjustment can be performed as a separate, second-stage analysis. AHRQ uses a two-stage approach to estimate its Quality Indicators, and the HCBS composites were adapted from a subset of the AHRQ Quality Indicators.
 - a. **Perform reliability adjustment on the estimated risk-adjusted rates.** Under this model, the estimated risk-adjusted rates from the first-stage model would serve as the outcome variable in the second-stage reliability-adjustment model.
 - b. **Perform reliability adjustment on the estimated state-level fixed effects.** If fixed effects for each state are included in the risk-adjustment model, their estimated values could also serve as the outcome variable in the second-stage reliability-adjustment model.

B. Specifying the distribution under which reliability adjustment is induced

The fundamental way in which reliability adjustment increases precision is by placing state-specific data in the broader context of data from other states. Statistically, this is carried out by assuming that states’ event rates are related to one another in that they come from a common underlying distribution. For this reason, an added complexity of the reliability-adjustment framework, regardless of whether we ultimately choose approach (1), (2a), or (2b) above, is that a statistical distribution must be chosen under which reliability adjustment is induced. In the Bayesian paradigm of statistical inference, such a distribution is called the “prior.”

A default choice for the prior distribution is a Normal (“bell-shaped”) distribution. The Normal distribution is used in the condition-specific mortality and readmission models used to assess hospital quality, as well as the AHRQ Quality Indicators. In addition to this default choice, we will also consider a Student’s t distribution, which has more probability density in its tails than the Normal and therefore induces less shrinkage of outlying values. For this reason, a Student’s t is often hailed as a robust alternative to the Normal distribution (Gelman and Hill 2007) in that it produces inference that more closely hews to the un-shrunken risk-adjusted estimates. Lastly, in the reliability-adjustment model of the risk-adjusted rates (approach 2a), we will also consider alternative distributions that may better reflect the underlying data-generating process (Wang et al. 2014) by explicitly accounting for the fact that these are count data.

A prior could be based on information contained only in the estimation sample and first-stage variables (in other words, the variables included in the risk-adjustment model). However, priors incorporating additional information, such as state characteristics or previous measure values, are also possible. We will consider including such information in the prior, for example 2005 HCBS composite estimates from AHRQ’s initial measure development report, state-level variation in PQI rates published by AHRQ, and for the 2010 HCBS user population, the 2009 HCBS state rates (Schultz et al. 2012; Torio and Andrews 2014).

C. Reliability-adjustment model validation

An important step in any statistical modeling activity is to assess the fit of the model to the data. This step is especially important for reliability-adjustment analyses, for which it is imperative to ensure that inference is not being unduly driven by model assumptions. To this end, we will conduct the most fundamental check of model fit (Gelman and Hill 2007; Gelman et al. 1996) by performing a “posterior predictive” analysis. To do this, we will first generate state-level HCBS composite rates under each candidate model. Such a replicated dataset resembles the data that might have been collected if the model were true. We will then compare the actual admission rates to each replicated dataset. Systematic differences between data and replications will provide an indication of poor model fit.

In addition, some candidate models may not have standard software available with which to fit them. In these cases, we will use Monte Carlo Markov Chain (MCMC) methods, which allow more flexibility than statistical software packages such as Stata and SAS. A cost of using MCMC methods, however, is that the statistician must ensure that the sampler, a numerical method, has converged to the true posterior distribution and that mixing has provided a sufficient number of effectively independent samples. We will perform additional diagnostics to assess MCMC performance.

V. COMPARISON OF STATE-LEVEL RESULTS

At the conclusion of this process, we will generate final risk- and reliability-adjusted results for the three HCBS composite measures for all available states, using 2009 and 2010 data. These results will begin to provide an indication of state-level differences in the quality of care delivered to HCBS recipients. However, determining how to identify “good” and “bad” quality is another crucial element for this work. We will explore several methods for interpreting state performance, as described in this section.

A. Standardization

The core of the risk- and reliability-adjustment model is a comparison of expected versus observed state-level performance, and we will specify the measures as indirectly standardized rates (rather than directly standardized rates) (Silber et al. 2014). An indirectly standardized rate will allow us to compare all HCBS recipients in a state to a common benchmark. The alternative—direct standardization—creates rates based on matched samples of HCBS recipients in each state. In line with the structure of the AHRQ Quality Indicators, indirect standardization has the benefit of including information on all HCBS recipients in a state, whereas the direct approach excludes persons with unique characteristics. The primary disadvantage of indirect standardization is that rates themselves are less useful for trying to understand why states’ HCBS composite rates vary. However, despite its limitations, indirect standardization lends itself to statistical comparisons to assess performance more easily than does direct standardization.

B. Common methods for evaluating performance

After producing the indirectly standardized state-level results, we will explore methods for assessing states’ performance relative to that of their peers. Table V.1 summarizes each approach, showing their distinct features and interpretation. The two most common approaches are to (1) rank the entity’s rate compared to that of its peers or (2) test the significance of the difference between the entity’s rate and a benchmark, where significance is measured using the confidence interval around the rate (Shwartz 2014).

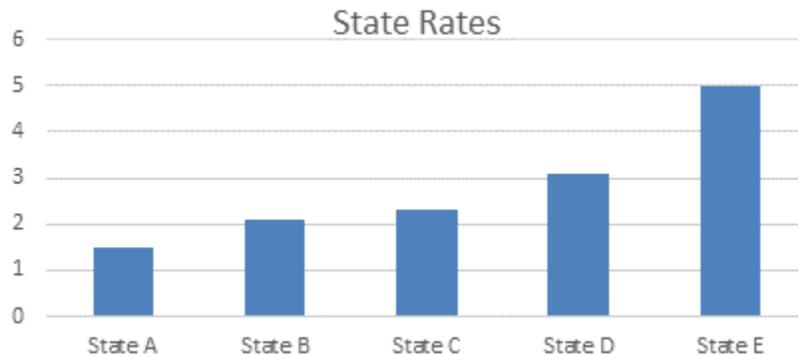
Table V.1. Summary of common methods for evaluating performance

Method	Description	Interpretation of lower rate ^a
Ranking	Ordering states based on their rates without making statistical inference	State A has the lowest rate, but this ranking may be due to chance
Performance Categorization	Distinguishing which states are statistically different from a benchmark without reference to the magnitude of the difference	There less than a 5% chance of observing such a low rate for State A if its true quality is no different from average
Exceedance Probability	Articulating the degree to which rates differ from a benchmark	State A has a 95% probability of being lower than the benchmark

^a This example is for interpreting results for a state with the lowest HCBS composite rate.

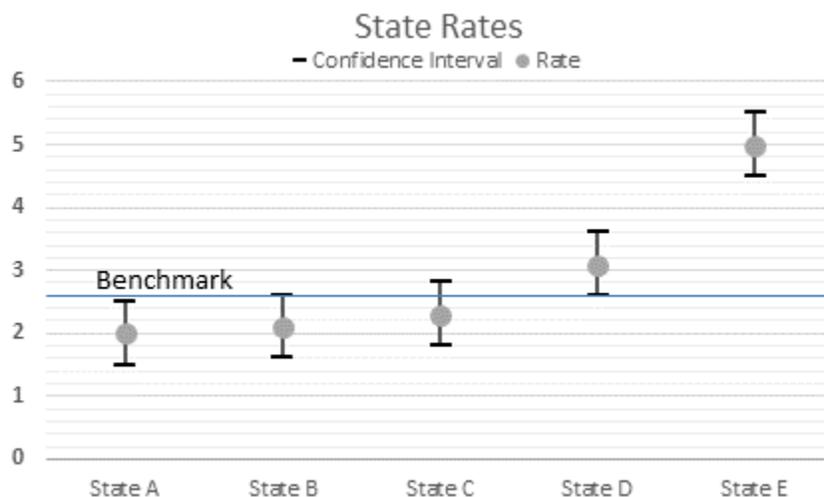
Ranking. In a simple ranking system, which was used to report results during the preliminary phase of this work, state-level results are ordered from lowest to highest, as shown in Figure V.1. This approach has the benefit of being straightforward, but does not permit us to assess the credibility of the state's rank and whether it is likely to reflect its performance or random variation.

Figure V.1. Simple ranking of state rates



Performance categorization. Another way to evaluate a state's performance is to compare its rate with a benchmark and test the hypothesis that the rate and benchmark differ. That benchmark may be defined as the norm among similar states, or all states, or some other appropriate peer group. The hypothesis is tested by determining whether the benchmark lies within the confidence interval around the state's rate. For example, if the goal of the policy is to identify the worst-performing states, it may be useful to determine which states have higher rates (worse performance) relative to the benchmark. Using this method, we would judge that State E is significantly higher (worse) on this measure because its confidence interval does not overlap with the benchmark (Figure V.2). However, that approach only permits us to assign a state's rate to categories of significantly higher, lower, or no different than or not significantly different from the benchmark. Particularly within the not significantly different category, a great deal of information about performance is discarded.

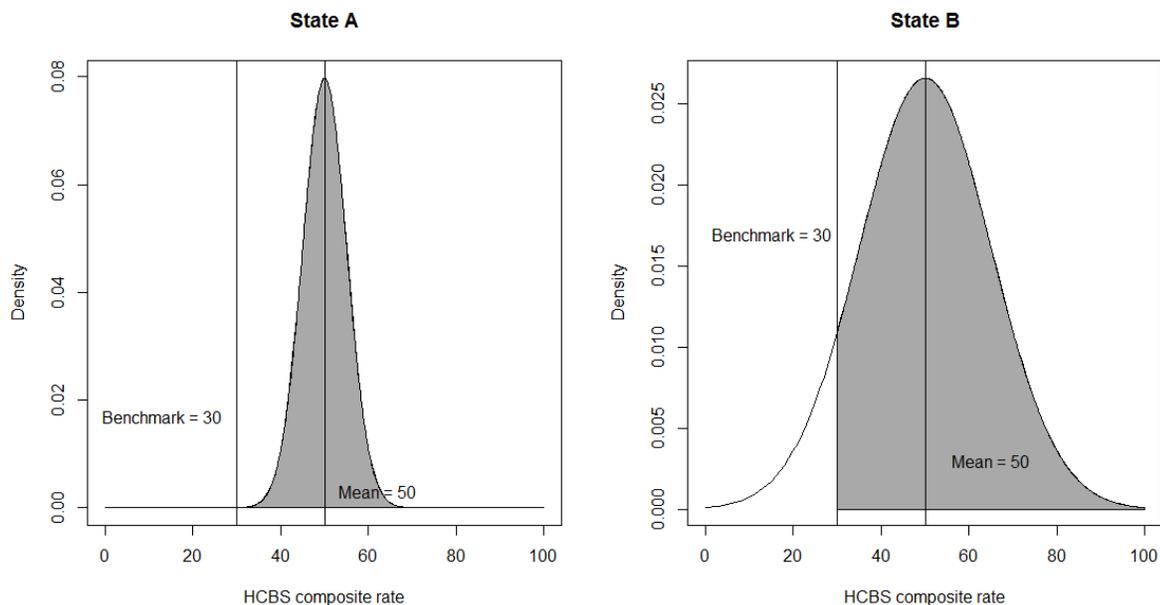
Figure V.2. State performance relative to a benchmark



Exceedance probability. The exceedance probability approach is an improvement upon the ranking and categorization methods described above in several regards. First, while simple ranking considers point estimates alone, exceedance probability conveys both the point estimate and the precision (that is, the variance or uncertainty of the estimated HCBS composite rate). Second, classifications based on confidence intervals do not provide information on the strength of the conclusion as to the states' performance relative to the benchmark, and exceedance probability does.

The exceedance probability approach is illustrated in Figure V.3. The estimated HCBS composite distribution is shown for two hypothetical states, A and B. The means of the estimated HCBS composite distributions are identical for the two states, and in this case, States A and B have mean rates that are greater than the benchmark. However, the variance of State A's estimated HCBS composite is less than the variance of the estimate for State B—in other words, the distribution for State A is less spread out than the distribution for State B. The exceedance probability is shown as the shaded gray area to the right of the benchmark. Although both states are above than the benchmark, the exceedance probability is higher for State A compared with State B.

Figure V.3. Example of exceedance probability



Notes: The gray shaded area shows the exceedance probability. States A and B have identical mean HCBS composite rates and are both above the benchmark. However, the distribution of State A's HCBS composite estimate is more variable than that of State B; therefore, the exceedance probability for State A is greater than that for State B.

C. Setting an appropriate benchmark

The intended goal of this effort is to provide states with information about the quality of services delivered to HCBS users in their state. As a result, it is the most useful to analyze state

performance relative to a set benchmark using an exceedance probability, as described above. Setting an appropriate benchmark is therefore a very important feature of this work, and we will explore a number of options in this regard.

Two potential options include using the national average or a decile cutoff, which are commonly employed in CMS hospital quality and payment programs. As a starting point, we will use the national average of each HCBS composite, and we will also develop peer-group or performance-relevant benchmarks.

In addition to these options, we will consider advanced methods described by Paddock (2014), including a histogram-based approach. This approach would simulate a posterior distribution of state-level HCBS composites using MCMC methods. These methods will help to adjust for any over- or under-dispersion in the HCBS composite distribution introduced through risk or reliability adjustment. The choice of benchmark will depend on the upstream model development, as well as input from the TEP.

It is conceivable that, based on TEP feedback, a variety of benchmarks may be developed. For example, national, peer-group (for example, neighboring state) or stratum-specific (for example, Medicaid-only beneficiary) benchmarks all provide different information that may be useful to states.

VI. APPLY THE MEASURES TO RELEVANT SUBPOPULATIONS

The end result of model development will include HCBS composite rates estimated for each state, a benchmark, and a framework for comparison. An additional objective of this effort is to develop a method for comparing state-level HCBS composite rates among a unique sub-population of HCBS users. For the purpose of model development, we focus on individuals who recently transitioned from institutionally based long-term care into HCBS through either the MFP program or other avenues. Because these populations are small compared to the general HCBS user populations, we anticipate first developing our risk- and reliability-adjustment approach for the overall HCBS user populations as described in the previous sections, and then making modifications as necessary. When applying the HCBS composites to small subpopulations, reliability adjustment becomes more important, because the estimated rates are more uncertain when less information (for example, sample size) is available to produce them. In the future, the HCBS composites could be implemented for other selected HCBS subpopulations, such as individuals with specific types of long-term needs or those enrolled in specific waiver programs.

Our recommendations on how to calculate these MFP and non-MFP state-level rates will build upon our team's extensive research on adjusting quality measures to support peer group comparisons. The recommendations will depend on (1) differences between the HCBS user population and the subpopulation, and (2) the sample size of the subpopulation. If the subpopulation is large and different from the HCBS population, it is best to completely recalibrate the HCBS composite to the subpopulation. However, if the subpopulation is small or similar to the HCBS population, it is possible to adjust the important parameters in the risk- and reliability-adjustment models to fit the subpopulation. We will address how to apply the HCBS composite measure to unique subpopulations through empirical analyses and through conversations with a TEP.

This page left blank for double-sided copying.

VII. NEXT STEPS

This report describes all of the key considerations and parameters in the measure development process that we will implement through August 2015. The contractual milestones over the remaining months include two TEPs and two reports (Table VII.1). The TEPs will provide external input on the measure development process, and the reports will document our progress and the finalized HCBS composite measure. The focus of the first TEP will be to provide feedback on candidate risk factors, and the technical approach for developing risk- and reliability-adjustment methods as outlined in this report.

The list of questions for the panel is a work in progress, but we expect the TEP to comment on the areas highlighted in Figure II.1 and possibly raise new questions for further TEP consideration. Issues pertaining to risk adjustment, reliability adjustment, and the comparison framework are higher priority items, while input on how to display the measure is a lower priority. In addition, if some topics require additional detail, we may propose post-meeting discussion with a subset of the TEP. The final TEP materials will prioritize our questions for the TEP, but below are some of the potential questions:

- **Information to include in the prior.** Reliability adjustment will improve the precision (reduce the variance) of HCBS composite estimates by incorporating information on our prior beliefs. The panel can help to weigh in on what information to include in the prior, such as the 2005 HCBS composites, information from the analytic sample, or information from neighboring or similar states (peer groups).
- **Stratification of rates by population.** Instead of producing HCBS composites for all HCBS recipients in a state, to produce results for specified HCBS user populations (for example, by MME status)?
- **Comparison framework.** Stakeholders like to make inferences based on quality measures, but quality measures are designed for certain uses and interpretations. Although the exceedance probability is relatively new to quality measurement, is it the best approach for meeting the HCBS composite needs?
- **Setting the benchmark.** With the goal of distinguishing performance between states, HCBS composite estimates will be compared to a benchmark. However, the expert panel will help us determine how to set appropriate benchmarks.
- **Displaying the results.** During the expert panel call on pressure ulcers, many comments alluded to the importance of the measure presentation or display. We will solicit input on what to display, what supporting technical information will aid interpretation, and what information related to hospitals' quality improvement efforts will contribute to their quality improvement processes.

Table VII.1. Key dates and deliverables

Date	Deliverable	Purpose
1/22/2015	HCBS Composite Methods Report	Report on refining the models, testing performance, and generating summary statistics
3/15/2015	TEP on HCBS Composite	Solicit feedback on new risk factors, modeling approach, and other technical issues
4/15/2015	HCBS Composite Preliminary Results Report	Finalized analytic approach, model performance, and state-level results
5/15/2015	Second HCBS TEP	Solicit feedback on preliminary modeling results
8/15/2015	Draft HCBS Risk-Adjustment Report	Produce a draft method for profiling states
9/17/2015	Final HCBS Risk-Adjustment Report	Finalize the method for profiling states

REFERENCES

- Agency for Healthcare Research and Quality. "Environmental Scan of Measures for Medicaid Title XIX Home and Community-Based Services Measure Scan as of July 5, 2007." Rockville, MD, Agency for Healthcare Research and Quality, June 2010. Available at <http://archive.ahrq.gov/professionals/systems/long-term-care/resources/hcbs/hcbsreport/hcbssum.html>
- Ash, A.S., S.E. Fienberg, T.A. Louis, S.L.T. Normand, T.A. Stukel, and J. Utts, with the Committee of Presidents of Statistical Societies. "Statistical Issues in Assessing Hospital Performance." Baltimore, MD: Centers for Medicare & Medicaid Services, 2011, revised 2012.
- Clark, T.S., and D.A. Linzer. "Should I Use Fixed or Random Effects?" *Political Science Research and Methods*, 2014, pp. 1–10.
- Davies, S.M., McDonald, K.M., Schmidt, E., Schultz, E., Geppert, J., and P.S. Romano. "Expanding Use of the Prevention Quality Indicators: Report of Clinical Expert Review Panel." Stanford, CA: Center for Primary Care and Outcomes Research, Stanford University, November 2009. Available at http://www.qualityindicators.ahrq.gov/Downloads/Modules/POI/POI_Summary_Report.pdf.
- Gelman, A., and J. Hill. *Data Analysis Using Regression and Multilevel/Hierarchical Models*. New York: Cambridge University Press, 2007.
- Gelman, A., X.L. Meng, and H. Stern. "Posterior Predictive Assessment of Model Fitness Via Realized Discrepancies." *Statistica Sinica*, vol. 6, no. 4, 1996, pp. 733–760.
- Long, J.S. *Regression Models for Categorical and Limited Dependent Variables*. Thousand Oaks, CA: Sage Publications, 1997.
- Osgood, D.W. "Poisson-based Regression Analysis of Aggregate Crime Rates." *Journal of Quantitative Criminology*, vol. 16, 2000, pp. 21–44.
- Paddock, S.M. "Statistical Benchmarks for Healthcare Provider Performance Assessment: A Comparison of Standard Approaches to a Hierarchical Bayesian Histogram-Based Method." *Health Services Research*, vol. 49, no. 3, 2014, pp. 1056–1073.
- Paternoster, R., and R. Brame. "Multiple Routes to Delinquency? A Test of Developmental and General Theories of Crime." *Criminology*, vol. 35, 1997, pp. 45–84.
- Ross, J. and A. Bohl. "Development and Testing of Risk-Adjusted Composite Measures for Populations Using Home- and Community-Based Services." Final report submitted to the Centers for Medicare & Medicaid Services. Cambridge, MA: Mathematica Policy Research, September 2013.

- Schultz, E., S. Davies, and K. McDonald. “Development of Quality Indicators for Home and Community-Based Services Population: Technical Report.” Stanford, CA: Center for Primary Care and Outcomes Research, Stanford University, June 2012. Available at http://www.qualityindicators.ahrq.gov/Downloads/Resources/Publications/2012/HCBS_QI_Technical_Report.pdf.
- Shwartz, M., E.A. Pekoz, J.F. Burgess, Jr., C.L. Christiansen, A.K. Rosen, and D. Berlowitz. “Probability Metric for Identifying High-Performing Facilities: An Application for Pay-for-Performance Programs.” *Medical Care*, vol. 52, no. 12, 2014, pp.1030–1036.
- Silber, J.H., P.R. Rosenbaum, R.N. Ross, J.M. Ludwig, W. Wang, B.A. Niknam, P.A. Saynisch, O. Even-Shoshan, R.R. Kelz, and L.A. Fleisher. “Hospital-Specific Template Matching for Auditing Hospital Cost and Quality.” *Health Services Research*, vol. 49, no. 5, 2014, pp. 1446–1474.
- Torio, C.M., and R.M. Andrews. “Geographic Variation in Potentially Preventable Hospitalizations for Acute and Chronic Conditions, 2005–2011.” HCUP Statistical Brief 178. Rockville, MD: Agency for Healthcare Research and Quality, September 2014. Available at <http://www.hcup-us.ahrq.gov/reports/statbriefs/sb178-Preventable-Hospitalizations-by-Region.pdf>.
- U. S. Congress. “Deficit Reduction Act of 2005.” Washington, DC: Government Printing Office, 2006. Available at <http://www.gpo.gov/fdsys/pkg/BILLS-109s1932enr/pdf/BILLS-109s1932enr.pdf>. Accessed November 19, 2012.
- Wang, S., M. Finucane, X. Lin, F. Yoon, and D. Jones. “Reliability Adjustment of the AHRQ Quality Indicators Under an Empirical Bayes Framework.” Submitted to the U.S. Department of Health and Human Services, Agency for Healthcare Research and Quality. Cambridge, MA: Mathematica Policy Research, December 2014.

APPENDIX A

METHODS

This page left blank for double-sided copying.

This section is taken directly from Ross and Bohl (2013). It describes the data and methods used to construct the HCBS user population and flag HCBS composite events.

METHODS

The risk-adjustment work proceeded in four phases. During the first phase, Mathematica identified the data sources and constructed the analytic files needed for this work. In the second phase, we worked with AHRQ to make necessary adjustments to the publicly available PQI software to identify qualifying ACSC events for Medicaid beneficiaries using HCBS. The third phase involved convening a group of HCBS researchers and risk-adjustment experts to provide feedback on our proposed risk factors and overall approach. The fourth and final phase encompassed the development and testing of our modeling approach and production of results for populations of interest. For model development, we utilized the population of Medicaid beneficiaries using HCBS in 2009, and risk-adjusted results were generated for the following groups:

1. Medicaid beneficiaries using HCBS in 2010
2. MFP participants who transitioned from 2008 to 2010
3. Medicaid beneficiaries who transitioned to HCBS outside of MFP from 2008 to 2010

A. Data Sources and Analytic File Construction

The analytic files developed for this work relied on both Medicare and Medicaid administrative data files because a majority of HCBS users are Medicare–Medicaid enrollees (MMEs). The analytic files required can be thought of in three primary segments: files for (1) the denominator population, (2) inpatient stays for individuals in the denominator, and (3) person-level chronic comorbidity information from the patient’s history used to establish patient risk. Unless otherwise indicated, the methods described in the following subsections apply to all four analytic populations used for this work.

Denominator Population

To establish the population of eligible Medicaid HCBS users, we used version 4.2 of the denominator specifications developed by AHRQ, which are summarized in Figure A.1. To first establish the baseline group of Medicaid beneficiaries (including Medicaid-only beneficiaries and MMEs) using HCBS services, we drew on the Medicaid administrative data from the Medicaid Analytic eXtract (MAX) data files, which contain person-level records of Medicaid enrollment and service use. To identify HCBS users, we used both indicators of enrollment in HCBS 1915(c) waivers from the MAX Person Summary (PS) file and receipt of HCBS under either a 1915(c) waiver or through the state plan from the MAX Other (OT) files.

Enrollment in an HCBS 1915(c) waiver was defined as at least one month of enrollment in the following waivers: aged/disabled, aged only, disabled only, traumatic brain injury, HIV/AIDS, mentally retarded/developmentally disabled, mental illness, technologically

dependent, an unspecified waiver, or autism.¹ Use of HCBS under a 1915(c) waiver was defined as at least one month of claims for personal care, at-home private duty nursing, adult day, home health of at least 90 days, residential care, at-home hospice, rehabilitation, case management, transportation, or durable medical equipment. Use of HCBS via the state plan was defined as at least one month of claims for personal care, at-home private duty nursing, adult day, home health of at least 90 days, residential care, or at-home hospice. The restriction requiring at least three consecutive months (90 days) of home health use is designed to eliminate those whose home health care is for rehabilitation purposes.

To identify the populations of MFP participants and Medicaid beneficiaries who transitioned to HCBS outside of MFP, we imposed some additional steps that were not relevant to the 2009 or 2010 HCBS user populations. For the MFP participants, we drew from Mathematica's MFP administrative files to identify participants between 2008 and 2010. Only MFP participants that had matching records in MAX PS files were retained. For the Medicaid beneficiaries who transitioned to HCBS outside of MFP between 2008 and 2010, we applied the methodology for identifying HCBS users described previously, but applied additional restrictions. For HCBS users between January 1, 2008 and March 31, 2010, a transition from institutional long-term care was identified if there were at least 181 days of care observed in an MFP-eligible institution (i.e., nursing facility, ICF-IID, or Mental Institution) and HCBS use occurred within three months of discharge. For HCBS users between April 2010 and December 2010, a transition from institutional long-term care was identified if there were at least 91 days of care observed in an MFP-eligible institution (i.e., nursing facility, institutional care facility for people with intellectual disabilities, institutions for mental diseases) and HCBS use occurred within three months of discharge.²

We imposed several important exclusions to all four populations, in accordance with the specifications developed by AHRQ. We excluded both Medicaid managed care and Medicare Advantage enrollees (Exclusion 1) because their claims were either not available or not comparable to those for beneficiaries in the fee-for-service system. To exclude Medicaid managed care enrollees, we used the MAX PS file to identify individuals enrolled in either a medical or comprehensive managed care plan, a long-term care managed care plan, or a Program of All-Inclusive Care for the Elderly program. In turn, we determined enrollment in Medicare Advantage using the monthly managed care flags available in the Medicare Beneficiary Summary File (MBSF). In both cases, if at least one month of managed care enrollment was identified during the period of interest, the individual was excluded from the measure denominator.

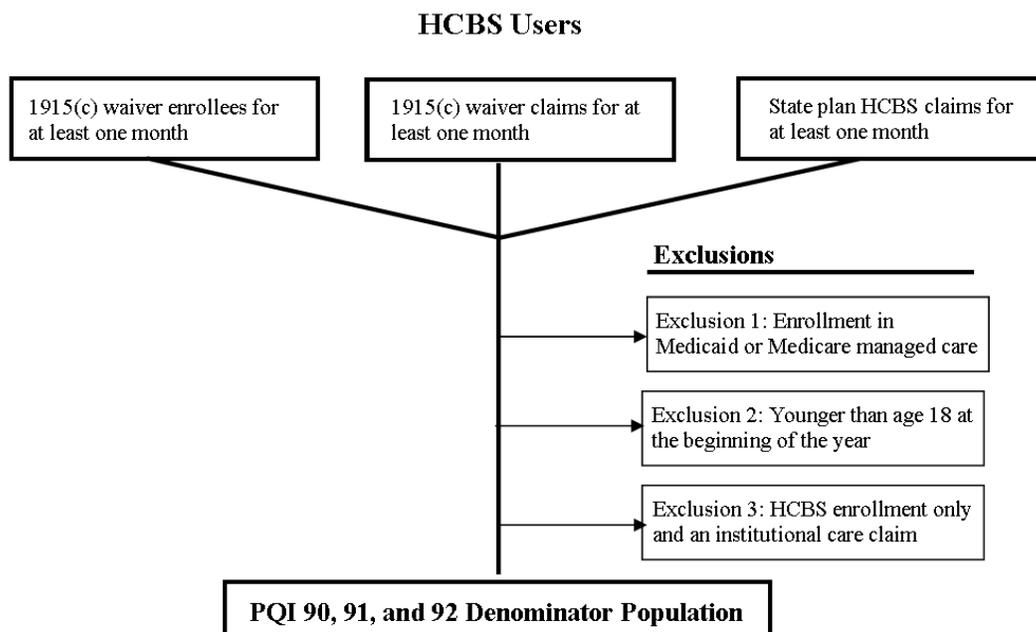
¹ The autism waiver was introduced after AHRQ's initial work to develop the HCBS measures was completed. Mathematica included these HCBS waivers in the standard definition of an HCBS user, however use of the autism waiver is generally restricted to children, who are not eligible for the measure denominators.

² This change was applied to ensure comparability with MFP program requirements. On March 23, 2010, as part of the Affordable Care Act, the criterion for MFP participation was reduced from a minimum of 180 days of institutional care to 90 days, not counting Medicare rehabilitation days. We applied this change as of April 1, 2010 for ease of data processing.

We also excluded children by limiting the eligible population to people 18 or older (Exclusion 2). We wanted to exclude this group because they are a population with substantially different care needs and propensity of ACSC events than adult HCBS users. Using the MAX PS file, we removed HCBS users who were younger than 18 at the start of the period of interest (for example, January 1, 2009, for the 2009 HCBS population).

Finally, people with a record of HCBS enrollment only (that is, no observed HCBS claims) who also had a record of institutional long-term care claims were excluded (Exclusion 3). We identified individuals who qualified as HCBS users only because they had at least one month of 1915(c) enrollment, but no 1915(c) or state plan HCBS claims, using the MAX PS and OT files, respectively. Then, if these individuals had at least one month with an institutional long-term care claim (that is, nursing home, intermediate care facilities for people with intellectual disabilities, or mental institution), they were excluded from the denominator population.

Figure A.1. Defining the Measure Denominator



Source: Adapted from Schultz E., S. Davies, and K. McDonald. “Development of Quality Indicators for Home and Community-Based Services Population: Technical Report.” Stanford, CA: Center for Primary Care and Outcomes Research, June 2012.

Inpatient Stays

Because the HCBS measures identify patient safety events that result in a hospitalization, a necessary step in our calculations required obtaining inpatient hospital stays for everyone in the eligible denominator population. Because a large proportion of Medicaid beneficiaries using HCBS is also eligible for and enrolled in Medicare (that is, they are MMEs), we drew on both Medicaid and Medicare data to identify inpatient hospital stays. For MMEs, we used the Medicare Provider Analysis and Review (MedPAR) and MAX Inpatient (IP) files to identify inpatient stays. For non-MME or Medicaid-only eligible individuals, we used MAX IP data only.

We determined MME status using the MAX PS file. Although an individual's MME status may change throughout a calendar year, we identified a beneficiary as an MME if they ever had that status during the year or other period of observation.

Candidate Risk Factors

The final step in our analytic file development was to add risk factors at the person level needed for the risk-adjustment models. We targeted the 27 Chronic Condition Warehouse (CCW) flags as our primary source for determining patient-level risk. For the MMEs in our sample, we obtained information on these 27 CCW flags from the MBSF³. The CCW flags are derived from Medicare administrative claims (for example, inpatient, outpatient, and pharmacy) and assessment (for example, the Minimum Data Set) records.⁴ During 2012, Buccaneer also produced these CCW flags for the Medicaid-only population. Unfortunately, first due to delays in obtaining permission to use these data and later delays in their availability on the CMS mainframe, we were not able to obtain the CCW flags for the non-MMEs in our sample populations in time for the initial model development work. As a result, the majority of report reviews model development and results for the MME population, although we do present preliminary findings for the combined MME and Medicaid-only population in Appendix C. Of note, for each population of interest identified in this report, the majority of our denominator population consisted of MME-eligible individuals.

B. Modifications to the AHRQ PQI Software

We used the PQI module from version 4.4 of the AHRQ Quality Indicator (QI) software to identify inpatient hospital admissions that resulted from an ACSC event.⁵ Version 4.4 was the most current version of the software at the time that this work began; since then, AHRQ has released version 4.5, which implements a change to the PQI 5 (COPD) measure, making the numerator more restrictive.⁶ We believe that this change will have only a minor impact on our results. We used the programs only in the PQI software that were necessary to identify the ACSC chronic, acute, and overall composite numerators, as well as the numerators for their component indicators (that is, the Control_PQI, PQFMTS, and PQSAS1 programs). We did not apply the basic age and sex risk-adjustment or area-level adjustment that is included in the software.

In addition to the selective use of the PQI software described earlier, we also implemented post-processing steps to identify the numerator events of interest. We implemented most of these

³ The 27 chronic condition flags are available for MMEs through the MBSF; they became available in July 2013 for non-MME beneficiaries.

⁴ Documentation on the Chronic Conditions Warehouse data can be found at <https://www.ccwdata.org/web/guest/home>. Accessed November 16, 2012.

⁵ The AHRQ PQI software is available at <http://www.qualityindicators.ahrq.gov/software/SAS.aspx>. Accessed November 16, 2012.

⁶ The log of *International Classification of Diseases, 9th edition, Clinical Modification* Coding Updates and Revisions to PQI Documentation and Software is available at http://www.qualityindicators.ahrq.gov/Downloads/Modules/PQI/V45/PQI_Changes_4.5.pdf. Accessed July 15, 2013.

to ensure that numerator events were not mistakenly double-counted. First, because MAX IP data do not contain a variable to indicate inpatient records representing a transfer, we identified inpatient hospital stays that had identical patient identifiers (Medicaid Statistical Information System [MSIS] ID and state), and in which the admission date matched the discharge date on another inpatient stay in our sample. We eliminated the second or transfer stay in this sequence; this is the “same-day readmission” approach described by AHRQ in the technical documentation for the HCBS QIs (Schultz et al. 2012). Second, we identified duplicate records in which patient identifiers, hospital, and admission and discharge dates were the same. In most of these cases, one record was from MedPAR and the other was from MAX IP data; we kept the MedPAR stay because MedPAR data are generally more complete than MAX IP data and, for MMEs, Medicare is the primary payer for inpatient care. Although AHRQ did not apply this second step during its original ACSC composite calculations, it was a recommendation in the final technical report (Schultz et al. 2012).

The final and most significant modification we applied was to limit numerator ACSC events to those that occurred during HCBS enrollment or use. To do this, we excluded numerator events for which we did not observe HCBS use (either via enrollment or claims, as described previously) either during the month before, the month of, or the month after the inpatient admission date. Limiting the numerator events in this way improves the attribution of an ACSC event to the quality of care delivered to an HCBS recipient.

www.mathematica-mpr.com

**Improving public well-being by conducting high quality,
objective research and data collection**

PRINCETON, NJ ■ ANN ARBOR, MI ■ CAMBRIDGE, MA ■ CHICAGO, IL ■ OAKLAND, CA ■ WASHINGTON, DC

MATHEMATICA
Policy Research

Mathematica® is a registered trademark
of Mathematica Policy Research, Inc.