

DATA QUALITY AND REPORTING RESOURCE 7: DATA VALIDATION PROCESS

Overview

Validating your data is important to ensure that each rate is an accurate, reliable reflection of the care that has been delivered and the outcomes a patient has experienced. Through the Population Health Management Initiative (PHMI), community health centers (CHCs) should have a validation process in place to ensure the accuracy of PHMI/HEDIS measurement.

This document delineates an approach for basic data validation and key steps to validate performance metric rates for the seven core HEDIS measures for PHMI. This process includes general steps (e.g., validating eligible population and stratified sub-populations), as well as steps specific to each core measure. Increased confidence in the accuracy of the rates reported will lead to better decisions regarding performance improvement interventions and ongoing monitoring.

For purposes of this document, data validation is the act of reviewing and confirming that the CHC data used to calculate the measures are of a minimum acceptable level of quality and accuracy. Depending on the analytics platform and process used to calculate each measure, validation can be an automated or a manual process or a combination.

Data validation can occur at each step in the calculation process of a measure to ensure the accuracy of that specific component.

Validation includes:

- Overall checks to verify the reasonability of the final rate (such as reasonability of the numerator and denominator).
- Specific checks to confirm that a service or visit was correctly coded on a member level.

The process presented here could also be applied to other quality measures, including PHMI supplemental measures, other non-PHMI measures or as part of a broader data governance program.

Data Validation Basics and Expectations Under PHMI Implementation

Reliable measurement is critical to the success of the PHMI implementation. Prior to reporting the core HEDIS measures for PHMI, the CHC should validate the numerator, denominator and overall eligible population for each measure.

Frequency of Validation and Confirmation

The level of rigor needed for ongoing data validation depends on successful validation and remediation of gaps in prior reporting periods and the level of automation or built-in validation in the analysis software/platform.

In general, CHCs who are using less automation should plan for a full validation of data with every submission. These CHCs could, over time, increase their levels of automated data and reduce the amount of manual validation. A CHC that uses an application to extract measurement data and has software with built-in certified data measures could expect to perform less validation with each submission.

Factors that will be considered in determining the frequency and level of data validation needed could include whether:

- There are no known gaps in reliability/accuracy of data from prior submissions/validations.
- The CHC is using an analytics software/platform with certified HEDIS measure source codes (see [certified vendor list](#)).¹
- The CHC is using an analytics software/platform with robust built-in validation checks.
- The CHC has coded processes for extraction of measurement data.
- The CHC has robust data governance processes and procedures and compliance with these processes.

CHCs not meeting the above criteria should conduct a rigorous data validation process of the core HEDIS measures for PHMI. With each submission of PHMI rates, CHCs will be asked to attest that they have validated their measure rates.

Passing Data Validation

A robust data validation process will allow CHCs to identify and remediate issues with data quality and reliability and ensure measure rates are an accurate reflection of patient care received and services provided. A minimum threshold for “passing” data validation standards is an impact of less than 5% percent on the rate. Impacts can be estimated based on the results of the data checks and primary source verification described below.

If data validation results indicate a known deviance greater than 5%, CHCs should:

1. Indicate data validation issues on their Data Reporting Tool (DRT) submission.
2. Develop and implement a remediation plan.

Data Validation Process Guidelines

As a best practice, CHCs should migrate from a manual validation process to a mostly automated process. Adopting a more automated approach offers a faster, more efficient and consistent way to extract the data for the core HEDIS measures for PHMI. It also offers improved data quality and integrity in other business processes and programs, like CalAIM or MCP P4P.

Even with robust software, some of the steps below would still apply to ensure the quality and accuracy of the CHC's core HEDIS measures for PHMI. Understanding where there are strong validation steps in place and where there are gaps will be critical to determining a process going forward.

In conjunction with their practice coach and the data quality and reporting subject matter experts (SMEs), the following is a general process that CHCs could follow to assess and improve their validation process:

Step 1: Analyze current state and identify software/methodology for producing each measure.

- Which software, if any, is the CHC currently using for each measure?
- Which version (i.e., are they using the current version)?
- Are built-in validation checks or processes available? If so, for what?

Step 2: Analyze overall data governance.

- What is the CHC's overall data governance structure?
- What policies and procedures are in place to ensure data governance?

- How is the integrity and validation of data ensured by the CHC's data governance structure?

Step 3: Determine gaps and level of data validation needed.

- Are there gaps in automated processes not covered by manual processes?
- Are there manual procedures that are not documented or consistently followed?
- Does the CHC lack a procedure to perform a data validation function?

Step 4: Explore gap remediation approaches.

- If not using one currently, is there an interest/ability in using an analytics option with certified measure reporting?
- A resource: [directory of vendors](#) that have earned measure certification.¹
- Identify equivalent manual processes to complete functions/fill gaps, including introducing policies and procedures as necessary.

Step 5: Develop sustainable processes for ongoing data integrity and validation.

- Ensure sustainable processes and develop policies and procedures as needed.

Population-Level Validation Checks

The clinic population is the basis from which all measures are derived. PHMI/HEDIS measures are reported for all Medi-Cal MCP assigned patients to the clinic. Ensuring that this population is accurate in the database or master file from which the measures are calculated is critical to ensuring the accuracy of the eligible populations for each measure and the segmented populations within each.

Each measure then requires validation of its unique data components, including numerators and denominators, to ensure the reliability and accuracy of the measure. Validation should consist of:

- 1) Hard checks for data discrepancies.
- 2) Reasonability checks for logical consistency.
- 3) Primary source verification.

Hard Checks

A hard check identifies discrepancies that are mathematically incongruent (e.g., a single race/ethnicity should not have a larger population than the total population).

Reasonability Checks

Reasonability checks determine whether the data is logical or consistent with other knowledge (e.g., a rate that is logical based on what is known about the patient population from other data metrics or sources).

When identifying the reasonability of populations, numerators and denominators, CHCs should consider what they know, such as:

- What is their knowledge of their patient population and do the numbers align?
- What other similar metrics does the CHC report (e.g., UDS) and do the core set measures (which utilize HEDIS specifications) seem reasonable in comparison?
- How do the CHC rates compare to regional state, or national benchmarks?
- How do the CHC rates compare to prior reporting periods?
- How do they compare to rates calculated by MCPs?

These questions will help to determine the “appearance” of reasonability. Some variation may be expected (for instance, UDS and HEDIS specifications differ), however, this process could detect differences the CHC would not expect to see based on the different specifications alone.

Detecting Expected Versus Unexpected Differences

Example 1: A CHC may know through other measures or trends that approximately 25% of the patient population has diabetes. If this prevalence is known, a CHC could determine that a denominator for the diabetes measure showing 5% or 95% of the health clinic population has diabetes does not appear reasonable.

Example 2: Compared to a prior reporting period of the same measure with no change in the specifications, neither the numerator nor the denominator should change significantly without another known change.

- A numerator can increase if process improvement activities are underway to improve the measure.
- A denominator can increase if there is a significant change in a CHC's population.

Numbers should not vary significantly without an explanatory factor.

Use the below table as a checklist to validate the overall population in the database or master file from which the measures are calculated, as well as the rates for each core HEDIS measure for PHMI.

FIGURE 7.1: VALIDATION CHECKLIST 1: HARD CHECKS AND REASONABILITY CHECKS

Validation Area	Validation Criteria	Y/N	Notes
Total Population Database/File Validation Steps			
Total Records	Number of records/patients in the reporting database/file matches the number extracted from the EHR/primary source.		
Total Eligible Population	Criteria used to identify the population applied with fidelity to PHMI/HEDIS specifications.		
	Population size appears reasonable and in alignment with other sources/knowledge.		
	Number of MCP-attributed patients aligns with the sum of patients on individual MCP-provided member attribution lists.		
	Sub-populations are not greater than the parent population.		
	Sub-populations (e.g., number of children, number of diabetics, race, and ethnicity) are in alignment with other sources/knowledge.		
Race/Ethnicity	Categorization of race and ethnicity applied with fidelity to PHMI specifications.		
	Numbers/percentages of patients by race and ethnicity appear reasonable and in alignment with other sources/ knowledge.		
	Number of unknown or unassigned patients appears reasonable and in alignment with other sources/knowledge.		
	Sum of patients delineated by race and ethnicity equals the total population.		
Site-Specific	Sum of patients delineated by clinic site equals the total population.		

FIGURE 7.1: VALIDATION CHECKLIST 1: HARD CHECKS AND REASONABILITY CHECKS (continued)

Validation Area	Validation Criteria	Y/N	Notes
Measure-Specific Validation Steps			
Hemoglobin A1c Control in Patients with Diabetes (Poor Control >9%)	Numerator appears reasonable and in alignment with other sources/knowledge.		
	Denominator appears reasonable and in alignment with other sources/knowledge.		
	Numerator is not greater than the denominator (note: the numerator can be equal to the denominator but not greater than it).		
	Segmented populations (e.g., race/ethnicity) align with expectations and are reasonable given the segmented population breakdown across health center.		
Controlling High Blood Pressure	Numerator appears reasonable and in alignment with other sources/knowledge.		
	Denominator appears reasonable and in alignment with other sources/knowledge.		
	Numerator is not greater than the denominator (note: the numerator can be equal to the denominator but not greater than it).		
	Segmented populations (e.g., race/ethnicity) align with expectations and are reasonable given the segmented population breakdown across health center.		
Prenatal and Postpartum Care (Postpartum)	Numerator appears reasonable and in alignment with other sources/knowledge.		
	Denominator appears reasonable and in alignment with other sources/knowledge.		
	Numerator is not greater than the denominator (note: the numerator can be equal to the denominator but not greater than it).		
	Segmented populations (e.g., race/ethnicity) align with expectations and are reasonable given the segmented population breakdown across health center.		

FIGURE 7.1: VALIDATION CHECKLIST 1: HARD CHECKS AND REASONABILITY CHECKS (continued)

Validation Area	Validation Criteria	Y/N	Notes
Colorectal Cancer Screening	Numerator appears reasonable and in alignment with other sources/knowledge.		
	Denominator appears reasonable and in alignment with other sources/knowledge.		
	Numerator is not greater than the denominator (note: the numerator can be equal to the denominator but not greater than it).		
	Segmented populations (e.g., race/ethnicity) align with expectations and are reasonable given the segmented population breakdown across health center.		
Well Child Visits in the First 30 Months of Life (First 15 Months)	Numerator appears reasonable and in alignment with other sources/knowledge.		
	Denominator appears reasonable and in alignment with other sources/knowledge.		
	Numerator is not greater than the denominator (note: the numerator can be equal to the denominator but not greater than it).		
	Segmented populations (e.g., race/ethnicity) align with expectations and are reasonable given the segmented population breakdown across health center.		
Child Immunization Status (Combo 10)	Numerator appears reasonable and in alignment with other sources/knowledge.		
	Denominator appears reasonable and in alignment with other sources/knowledge.		
	Numerator is not greater than the denominator (note: the numerator can be equal to the denominator but not greater than it).		
	Segmented populations (e.g., race/ethnicity) align with expectations and are reasonable given the segmented population breakdown across health center.		

FIGURE 7.1: VALIDATION CHECKLIST 1: HARD CHECKS AND REASONABILITY CHECKS (continued)

Validation Area	Validation Criteria	Y/N	Notes
Depression Screening and Follow-Up for Adolescents and Adults	Numerator 1 (patients screened) appears reasonable and in alignment with other sources/knowledge.		
	Denominator 1 (patients 12+ years of age) appears reasonable and in alignment with other sources/knowledge.		
	Numerator 2 (patients followed up) appears reasonable and in alignment with other sources/knowledge.		
	Denominator 2 (patients screened positive) appears reasonable and in alignment with other sources/knowledge.		
	Numerators are not greater than the denominators (note: the numerator can be equal to the denominator but not greater than it).		
	Segmented populations (e.g., race/ethnicity) align with expectations and are reasonable given the segmented population breakdown across health center.		

Primary Source Verification

Primary source verification (PSV) is a common best practice. For example, health plans reporting HEDIS measures must go through PSV with a HEDIS auditor for certain types of data to ensure its accuracy and reliability for use in reporting.

PSV involves:

- Using the patient-level data file for a measure. This is a file that identifies all patients in the denominator with a flag indicating whether they were numerator-compliant.
- Tracing that patient back to their primary source of data (e.g., the medical record) to ensure that the documentation supports the patient having been included in the measure and whether they are compliant.

An Example of Primary Source Verification

A patient who was identified as numerator-compliant for the colorectal cancer screening measure should have evidence in the medical record of factors that comply with the measure. This would include:

- Patient is 45 to 75 years of age.
- Patient has documented evidence of a colorectal cancer screening within the time frame allowed for the particular type of colon cancer screening received.

Recommended Approach to Primary Source Verification

PSV can be a resource-intensive process. For PHMI, initial PSV could consider a random sample of five patients from each measure. If the primary source verified the information contained in the patient-level file for those patients, the measure would pass PSV. If discrepancies were detected, the CHC should:

- Select a larger sample (up to an additional 45 records, as a best practice).
- Assess results of the larger sample to determine if deficiencies are detected.
 - » If no, PSV is complete.
 - » If yes, the CHC should identify gaps and determine whether the gaps are isolated or pervasive.

Based on findings, the CHC should develop an appropriate remediation strategy and report data validation issues on its data reporting tool (DRT) submission. CHCs should use PSV to validate their PHMI/HEDIS rates:

- As an initial process and continuing with each submission until no issues are found during PSV.
- Any time there is a material change in how the measures are pulled or data sources used.

Use Figure 7.2: Validation Checklist 2: Primary Source Verification below as a checklist when conducting primary source verification of the overall population in the database or master file from which the measures are calculated, as well as rates for each core HEDIS measure for PHMI.

FIGURE 7.2: VALIDATION CHECKLIST 2: PRIMARY SOURCE VERIFICATION

Validation Area	Validation Criteria	Y/N	Notes
Total Population Database/File PSV			
Race/Ethnicity	Race/ethnicity of patient in file aligns with race/ethnicity in patient medical record.		
MCP-Attributed Patients	Patients can be traced back to MCP-provided membership files.		
Measure-Specific PSV			
Hemoglobin A1c Control in Patients with Diabetes (Poor Control >9%)	Diabetes diagnosis.		
	HbA1c value missing or value >9%.		
Controlling High Blood Pressure	Two HTN diagnoses.		
	Latest BP reading <140/90 mm Hg.		
	Latest BP reading is after second HTN diagnosis.		
Prenatal and Postpartum Care (Postpartum)	Delivery date between October 8 of the previous year and October 7 of the measurement year.		
	Postpartum visit within seven to 84 days of delivery date.		
Colorectal Cancer Screening	Aged 45 to 75 years.		
	Colorectal cancer screening and date (within range based on type of screening): 1. Fecal occult blood test (within the year). 2. Stool DNA (sDNA) with FIT test (within past three years). 3. Flexible sigmoidoscopy (within past five years). 4. CT colonography (within past five years). 5. Colonoscopy (within the past 10 years).		
Well Child Visits in the First 30 Months of Life (First 15 Months)	Patient turned 15 months old in measurement year.		
	Dates for six or more well child visits (or another visit with all the components of a well child check documented).		

FIGURE 7.2: VALIDATION CHECKLIST 2: PRIMARY SOURCE VERIFICATION (continued)

Validation Area	Validation Criteria	Y/N	Notes
Child Immunization Status (Combo 10)	Patient turned two years old in the measurement year.		
	Patient has all 10 applicable immunizations: 1. 4 DTAP (diphtheria, tetanus, acellular pertussis). 2. 3 IPV (polio). 3. 1 MMR (measles, mumps, rubella). 4. 3 HIB (haemophilus influenza type B). 5. 3 HEP B (hepatitis B). 6. 1 VZV (chicken pox). 7. 4 PCV (pneumococcal conjugate). 8. 1 HEP A (hepatitis A). 9. 2 or 3 RV (rotavirus—2 Rotarix; 3 Rota Teq). 10. 2 Influenza (flu).		
Depression Screening and Follow-Up for Adolescents and Adults	Patient is 12+ years old.		
	Diagnosis of depression.		
	Screening with a standardized instrument: 1. Patient Health Questionnaire (PHQ-9, PHQ-9M, PHQ-2). 2. Beck Depression Inventory (BDI-II). adults only 3. Beck Depression Inventory-Fast Screen (BDI-FS). 4. Center for Epidemiologic Studies Depression Scale-Revised (CESD-R). 5. Edinburgh Postnatal Depression Scale (EPDS). 6. PROMIS Depression. 7. Duke Anxiety-Depression Scale (DUKE-AD). adults only 8. Geriatric Depression Scale—Short Form and Long Form (GDS). adults only 9. My Mood Monitor (M-3). adults only 10. Clinically Useful Depression Outcome Scale (CUDOS). adults only		
	Positive result on screening.		
	Follow-up within 30 days of screening.		

ENDNOTES

- 1 National Committee for Quality Assurance. HEDIS® and AMP Vendor Certification Status; [July 10, 2023]. Available from: https://www.ncqa.org/wp-content/uploads/2022/11/MY2023_MeasureCertification_VendorList.pdf.