OMB Approval No.
0938-0786

EQR PROTOCOL 3: VALIDATING PERFORMANCE IMPROVEMENT PROJECTS (PIPs)

A Mandatory Protocol for External Quality Reviews (EQR)

Protocol 1: Assessment of Compliance with Medicaid Managed Care Regulations
Protocol 2: Validation of Measures Reported by the MCO
Protocol 3: Validation of Performance Improvement Projects (PIPs)
Protocol 4: Validation of Encounter Data Reported by the MCO
Protocol 5: Validation and Implementation of Surveys
Protocol 6: Calculation of Performance Measures
Protocol 7: Implementation of Performance Improvement Projects (PIPs)
Protocol 8: Focused Studies
Appendix V: Information Systems Capabilities Assessment (ISCA)

Department of Health and Human Services (HHS)
Centers for Medicare & Medicaid Services (CMS)

Protocol 3
Version 2.0

December 1, 2011

Table of Contents

PURPOSE AND OVERVIEW OF THE PROTOCOL

This mandatory protocol is used to determine whether a health care quality performance improvement project (PIP) was designed, conducted, and reported in a methodologically sound manner. The purpose of a PIP is to assess and improve the processes and outcomes of health care provided by an MCO. Protocol 3 specifies procedures for EQROs to use in assessing the validity and reliability of a PIP.[1] Protocol 3 specifies how to conduct the following three activities:

1. Assess the study methodology;
2. Verify PIP study findings; and
3. Evaluate overall validity and reliability of study results.

Results of the MCO's PIPs may be reported in the annual Secretary's Report on the Quality of Care for Children in Medicaid and CHIP or the annual Secretary's Report on the Quality of Care for Adults in Medicaid. These reports are released every September and information that is not available from a State's EQR report may be so noted in the reports. Both reports will be available on the CMS Medicaid website. States are strongly encouraged to have EQROs include PIP outcome and trending information reported in the EQR technical report. This will enable the Secretary to include results and lessons learned from State intervention strategies to improve care as part of that annual reporting process.

Additionally, States may incorporate specific PIPs as part of their State quality strategy, required by Section 1932(c)(1) of the Social Security Act, to align with the HHS National Quality Strategy for Quality Improvement in Health Care. When doing so, soliciting input from participating MCOs/PIHPs in the identification of PIP topics and methodologies may be helpful so that relevant clinical, administrative and population-based improvement efforts are addressed as part of the State's overall strategy to improve health care delivery and outcomes of the people it serves.

ACTIVITY 1: ASSESS THE STUDY METHODOLOGY

Activity 1 includes reviewing the following steps:

1. Review the selected study topic(s);
2. Review the study question(s);
3. Review the selected study indicators;
4. Review the identified study population;
5. Review sampling methods (if sampling used);
6. Review the data collection procedures;
7. Assess the MCO's Improvement strategies;
8. Review the data analysis and interpretation of study results;

[1] This protocol relies heavily on a guidebook produced by the National Committee for Quality Assurance (NCQA) that identifies key concepts in quality improvement (QI) studies. Please see References at the end of this protocol for a list of references that were used to develop this protocol.

9. Assess the likelihood that reported improvement is “real” improvement; and
10. Assess the sustainability of documented improvement.

The EQRO will review the PIP design and implementation using documents provided by the MCO, which may be supplemented with MCO staff interviews. In addition, the MCO, on an ad hoc basis, may supplement information obtained through hardcopy, electronic submission, or interviews.

The EQRO should follow the 10 steps below and answer the questions posed in each. The answers should be recorded on a standardized PIP Validation Worksheet (see Attachment A).

Step 1: Review the Selected Study Topic(s)

In this step, the reviewer determines the appropriateness of the selected study topic(s). The topic(s) should address the overarching goal of a PIP, which is to improve processes and outcomes of health care provided by the MCO.

Criteria

The PIP should target improvement in either clinical or non-clinical services delivered by the MCO. Study topics must reflect MCO enrollee characteristics including demographics, prevalence of disease, and the potential consequences of disease. The project may focus on patterns of over or under utilization that present a clear threat to health or functional status. The State may select the MCO’s study topic(s). Topics may also be selected based on enrollee input.

The topic should address a significant portion of the enrollees (or a specified sub-portion of enrollees) and have the potential to significantly impact enrollee health, functional status, or satisfaction. The topics should reflect high-volume or high-risk conditions of the population served. High-risk conditions may occur for infrequent conditions or services. High risk also exists for populations with special health care needs, such as children in foster care, adults with disabilities, and the homeless. Although these individuals may be small in number, their special health care needs place them at high risk.

The CMS suggests that States consider PIPs which address some of the national health priorities CMS has identified, (e.g., in 2011, Partnership for Patients, Million Hearts Campaign, pediatric oral health, and childhood obesity).

Recommended Sources of Supporting Information

- Data about enrollees:
    - Health risks;
    - Utilization of clinical or non-clinical services;
    - Demographics (Age, sex, race, ethnicity, language); and
    - Disability or functional status
    - Geographic location of membership
- Utilization, diagnostic, and outcome information on:

    - Outpatient and inpatient encounters, services, and procedures;
    - Medications and devices;
    - Diagnoses; and
    - Adverse incidents (such as deaths, avoidable admissions, or readmissions)
- Standardized local, State, or national measures when appropriate and available
- Data from other outside organizations, such as Medicaid or Medicare fee-for-service data; data from other health plans; and local or national public health reports on conditions or risks for specified populations; data from health information exchange technology – including registries.
- Data from surveys, grievance and appeals processes, disenrollment, and requests to change providers
- Data on appointments and provider networks (e.g., access, open and closed panels, and provider language spoken)
- Data from certified electronic health record (EHR) technology as described in Appendix V: Information System Capabilities Assessment
- Data from previous EQRO focused surveys

Assessment
To determine appropriateness of the study topic:

A. Review the documentation justifying the study topic using the potential data sources listed above.
    1. Did the State require the PIP topic, goal and/or study methodology?
    2. Were specific MCO or State enrollee demographic characteristics and health risks considered?
    3. Is the topic consistent with demographic and epidemiologic information of the current enrollees?

B. Did the MCO consider input from enrollees who are users of, or concerned with specific areas such as mental health or substance abuse?

C. The PIP, over time, should address a broad spectrum of enrollee care and services. Does the PIP address:
    1. Children with special health care needs?
    2. Preventive care?
    3. Acute and chronic condition care?
    4. High-volume and high-risk services (even if they are low frequency)?
    5. Specialized care received from centers such as burn, transplant, and cardiac surgery centers?
    6. Continuity or coordination of care when received from multiple providers and multiple episodes of care?
    7. Appeals and grievances?
    8. Access to and availability of care?

Step 2: Review the Study Question(s)

In this step, the reviewer determines the appropriateness and adequacy of the study (questions). The study question(s) identifies the focus of the PIP and establishes the framework for data collection, analysis, and interpretation.

Criteria
The study question(s) should be clear, simple, and answerable. In addition, they should be stated in a way that supports the ability to determine whether the intervention has a measurable impact for a clearly defined population.

An example of a vague study question is:

- ✗ "Does the MCO adequately address psychological problems in patients recovering from myocardial infarction?"

  In this example, it is not clear how "adequately address" will be assessed. Furthermore, "psychological factors" is not specific.

A clearer study question is:

- ✓ "Does the intervention reduce the likelihood that patients with myocardial infarction will develop severe emotional depression during hospitalization?"

Potential Sources of Supporting Information:

- QI study documentation
- Relevant clinical literature
- Enrollee focus groups/surveys
- Enrollee/provider representatives on Quality Committees

Step 3: Review the Identified Study Population

Measurement and improvement efforts must be system-wide.

Criteria
The PIP must clearly identify the 'system' or population, also referred to as the universe. Once the population is identified, the MCO will determine whether to study data for the entire population or a sample of that population. A representative sample of the identified population is acceptable (see Step 5).

Potential Sources of Supporting Information
Data on the MCO's enrolled population as well as enrollee counts relevant to the study topic and measures. This includes:

- Demographic information from the MCO's enrollment files
- The MCO's utilization and outcome information such as:

- Services
- Procedures
- Admitting and encounter diagnoses
- Adverse incidents (e.g., deaths, avoidable admissions, readmissions)
- Patterns of referrals
- Authorization requests

- Other databases, as needed (e.g., pharmacy claims data to identify patients taking a specific medication(s) during a specific enrollment period).

Assessment
Review the study description and methodology to determine if the study clearly identified the study population. Consider the following questions:

A. How was the "at risk" population defined?
B. Are all individuals clearly defined in terms of the identified study question(s) and relevant indicators?
C. Is the entire study population or a sample used? If the organization is able to collect and analyze data through an automated data system, it is possible to study the whole population? If the data must be collected manually, sampling may be more realistic.
D. Did the definition of the study population include requirements for the length of the study populations' members' enrollment in the MCO? The required length of time will vary depending on the study topic and study indicators.
E. If the entire population was studied, did the data collection approach capture all enrollees to whom the study question applied?
F. If a sample was used, go to Step 5. If the entire population was studied, skip Step 5 and go to Step 6. If HEDIS® measures and sampling methodology is used, go to Step 7.

Step 4: Review the Selected Study Indicators

A study indicator is a measurable characteristic, quality, trait, or attribute of a particular individual, object, or situation being studied. Indicators may be quantitative or qualitative and continuous or discrete. Discrete or categorical indicators have a limited number of possible categories (e.g., an individual has/has not received a flu shot in the last 12 months). In contrast, continuous indicators have unlimited possible values within the limits the indicator range, (e.g., age, blood pressure, temperature). Data collected on a continuous indicator such as blood pressure can be used for a discrete indicator, (e.g., an enrollee's blood pressure is/is not below a specified level).

Criteria
Each PIP should have one or more measured indicator to track performance and improvement over a specific period of time. All measured indicators should be:

- Objective; and

- Clearly defined; and
- Based on current clinical knowledge or health services research; and
- Enrollee outcomes (e.g., health or functional status, enrollee satisfaction); or
- A valid indicator of these outcomes.

The number and complexity of measures may vary depending on:

- The study question(s);
- The complexity of existing practice guidelines for a clinical condition; and
- Availability of data and resources to gather data

Potential Sources of Supporting Information

- Clinical and non-clinical practice guidelines
- Administrative data
- Medical records

Assessment

The EQRO will review the project documentation to determine if appropriate measures are used. Examples of measures currently existing within the public health community or the managed care industry include NCQA's Healthcare Effectiveness Data Information Set (HEDIS®) or measures that are developed by CMS and AHRQ (such as the Pediatric or Adult Core Measures). The MCO may also develop measures based on current clinical practice guidelines or health services research. When an MCO develops its own measures, it must document the basis for its adoption. Consider the following questions:

A. Did the study use objective, clearly defined, measurable, time-specific indicators?

B. Do the measures capture changes in health status, functional status, or enrollee satisfaction?

C. Do the measures have any of the following key characteristics:
    - related to identified health care guidelines relevant to the study question?
    - an important aspect of care or operations that made a difference to the MCO's/ beneficiaries?
    - data available through administrative, medical records or another readily accessible source?
    - limitations on the ability to collect the data skew the results?
    - require explicit or implicit criteria (Note that the specificity of the criteria used to determine compliance with a measure must be considered)?
    - if relevant, a strategy to ensure inter-rater reliability?

Notes to Reviewers

For the purpose of this protocol, "outcomes" are defined as changes in patient health, functional status, or satisfaction resulting from the PIP. For a PIP with a clinical focus, measures should

include change in health status or functional status or process of care proxies for these outcomes. Standardized performance measures addressing outcomes may be limited because health outcomes are influenced by factors outside of the organization's control, such as poverty, genetics, and environment. For these reasons, quality measures do not always need to be outcome measures.

Process measures, while acceptable, must offer strong clinical evidence that the process being measured is meaningfully associated with outcomes. This determination should be based on published guidelines, including citations from randomized clinical trials, case control studies, or cohort studies. At a minimum, the PIP should be able to demonstrate a consensus among relevant practitioners with expertise in the defined area who attest to the importance of a given process. It will be important that MCOs note within their PIP external validity threats which could affect the results of the outcome measures.

While enrollee satisfaction is an important outcome of care in clinical areas, improvement in satisfaction should not be the only measured outcome of a clinical project. Some improvement in health or functional status should be addressed. For projects in non-clinical areas, the use of health or functional status measures is preferred, but not required, when addressing access or availability of services. Enrollee satisfaction alone may be sufficient for some non-clinical projects.

Step 5: Review Sampling Methods

In this step, the reviewer determines the appropriateness and validity of the PIP's sampling methods. A sample is a statistical subset of a population that represents the entire population. There are several types of sampling methods that are appropriate for different types of PIPs.

Criteria

Appropriate sampling is necessary to ensure valid and reliable information. Please refer to Appendix II of the EQR Protocols for an overview of sampling methodologies applicable to PIPs. MCOs that use HEDIS® measures should also use HEDIS® sampling methodology, which is considered valid and reliable.

Potential Sources of Supporting Information

Data on enrollee characteristics relevant to health risks or utilization of clinical and non-clinical services including:

- Age;
- Sex;
- Race/ethnicity;
- Language; and
- Functional status.

Utilization information includes:

- Diagnostic and outcome information such as:
  - Procedures,
  - Admitting and encounter diagnoses,
  - Adverse incidents (such as deaths, avoidable admissions, readmissions),
  - Patterns of referrals, and
  - Authorization requests;
- Other information as needed, such as pharmacy claims data to identify patients taking a defined number of a specific medication(s) during a specific enrollment period.

Assessment
Review the study description and methodology. Consider following questions:

A. Did the methods
    1. Calculate the required sample size?
    2. Consider and specify the true or estimated frequency of the event?
    3. Identify the confidence level to be used?
    4. Identify an acceptable margin of error?

B. Are valid sampling techniques used?

Sampling
Statistical sampling methods apply two basic methodologies- probability sampling and non-probability sampling. General information about using various types of sampling methods is provided in Appendix II of the EQR Protocols.

Probability sampling is also known as random sampling, which means leaving the selection of population units totally to chance and removing biased selection of study subjects. An example would be a study of how many women received a cervical cancer screening during a specified year by randomly selecting 100 of the 1,000 women members of the MCO. Types of probability sampling include:

- Simple Random Sampling;
- Systematic Random Sampling;
- Stratified Random Sampling; and
- Cluster Sampling.

Non-Probability sampling uses specific characteristics of the study subject. An example would be a study of the performance of a group practice by sampling all the patients that were seen in that office on a specific day. Types of non-probability sampling include:

- Judgment Sampling;
- Convenience Sampling; and
- Quota Sampling.

Step 6: Review the Data Collection Procedures

In this step, the reviewer determines the validity of the procedures the MCO uses to collect the data that inform the PIP measurements. Study results are dependent on accurate and valid data that are collected appropriately.

Criteria
Data collection procedures must ensure that the data used to measure performance are valid and reliable. Valid data measure what is intended to be measured, while reliable data produces consistent results. To ensure both validity and reliability, the data collection plan should specify:

- The data to be collected;
- The data sources;
- How and when the data are to be collected;
- Who will collect the data; and
- Instruments used to collect the data.

To ensure the collection of valid and reliable data, the MCO should develop collection specifications appropriate to the type of data needed. Procedures for collecting data from automated data systems will be different from procedures for visual inspection of medical records or other primary source documents. However, both types of data collection require the following to ensure the data are consistently extracted and recorded:

- Qualified Personnel:
  Data collection personnel have the conceptual and organizational skills to abstract the data. The specific skills will vary depending on the nature of the data and the degree of professional judgment required. For example, experienced clinical staff, such as registered nurses, should be used to extract the appropriate data from medical records to support a judgment about whether clinical criteria are met. In contrast, trained medical assistants or medical records clerks may collect data if the abstraction involves verifying the presence of a diagnostic test report.

- Inter-Rater Reliability:
  The number of data collection staff used for a given project affects the reliability of the data. A smaller number of staff promotes inter-rater reliability; however, it may also increase the amount of time it takes to complete this task. The PIP should also consider and address intra-rater reliability (i.e., reproducibility of judgments by the same abstractor at a different time).

- Guidelines for Obtaining and Recording the Data to ensure consistent interpretation among and between data collection staff. This is particularly important when there are multiple reviewers collecting data. Appropriately qualified data collection staff (e.g., registered nurse, certified coder, etc.) should have access to a glossary of terms for each project before data collection begins. The data collection staff should be provided with

clear and succinct written instructions, including an overview of the study, specific instructions on how to complete each section of the form or instrument, and general guidance on how to handle situations not covered by the instructions.

Potential Sources of Supporting Information

- List of sources of data used in the study.
- If medical record review or other manual data collection is used to produce study data:
  - data recording forms; and
  - instructions to data collectors.
- If automated data collection is used, an algorithm showing the steps in the production of quality indicators and other relevant data collection.
- When assessing non-clinical services such as health care access or cultural competency or care coordination, a study may utilize information on how the MCO is structured and operates.

Assessment

Two processes may be used to assess data collection procedures:

1. Reviewing the study's approach to data collection (discussed in this step); and
2. Conducting a verification sample of the study's findings (discussed in Activity 2 of this protocol).

Consider the following questions to determine the appropriateness of the PIP's data collection procedures:

A. Does the study design clearly specify how the data are to be collected?

   Accurate measurement depends on clear and concise definitions of data elements. When descriptive terms are used (e.g., high, low, or normal), numerical definitions must be established for each term. The units of measure (e.g., pounds, kilograms, etc.) must also be specified.

B. Does the study design clearly specify the sources of data?

   Data sources vary and depend upon the selected topic and indicators. The topic and indicators will reflect clinical and research considerations and the available MCO data sources. Sources can include:

   - Beneficiary medical records;
   - Tracking logs;
   - Encounter and claims systems;
   - Provider interviews;
   - Beneficiary interviews; and
   - Surveys.

C. Does the study design specify a systematic method of collecting valid and reliable data that represents the entire population relevant to the study (sampling adequacy)?

D. What is the type of data collected (automated vs. manual)?

Automated Data Collection: Evaluating an automated data collection methodology emphasizes the system that stores the data and should focus on an estimation of the degree of completeness of the automated data used for the PIP study indicators.[2] For example:

- Inpatient data: Did the data system capture all inpatient admissions?
- Primary care data: Did primary care providers submit encounter data for all encounters?
- Specialty care data: Did specialty care providers submit encounter data for all encounters?
- Ancillary services data: Did ancillary service providers submit encounter or utilization data for all services provided?
- EHR data: Was patient clinical, service, or quality metrics data retrieved from certified electronic health record technology?

Manual Data Collection: This may be the only feasible option for MCOs and selected topics and emphasizes who and how the data are abstracted. The beneficiary medical record is the most frequently used data source. Other manual systems include clinical tracking logs, registries, complaint logs, and manual claims. When evaluating manual data collection, consider the following:

- Is qualified staff collecting the data?
- Does the staff have the requisite clinical knowledge and skills, including good conceptual skills, organization skills, thoroughness, and strong documentation skills?
- Does the data collection tool provide reliable and accurate data collection over the time periods studied?
- Is the data collection instrument(s) used for manual data collection clear and promote inter-rater reliability?

E. Does the study design prospectively specify a data analysis plan that reflects the type of data being collected (i.e., qualitative, quantitative data, or both)?

Qualitative data describes characteristics or attributes by which persons or things can be classified (e.g., sex, race, poverty level, or the presence or absence of a specific disease). Calculation of proportions and calculation of rates are the two most common qualitative measures.

[2]The accuracy of automated data is also a concern, but validation of this is beyond the scope of this protocol.

Quantitative data are concerned with numerical indicators such as height, weight and blood levels. The methods by which the data are analyzed and presented will vary by type of data. Quantitative data require, at a minimum, simple descriptive statistics such as measures of central tendency (i.e., mean, median, or mode) and measure of variability (i.e., range or standard deviation).

F. Are data collected on the entire population or a sample?

G. Is the PIP comparing these results to those of previous or similar studies? If so, the data analysis plan should evaluate the comparability of the studies and identify the appropriate statistical tests.

H. Is the PIP comparing the performance of an individual MCO, a number of MCOs, or different provider sites? Comparing the performance of multiple entities involves greater statistical design and analytical considerations than those required for a study of a single entity, such as a MCO.

Step 7: Review Data Analysis and Interpretation of Study Results

In this step, the reviewer determines the accuracy of the MCO's plan for analyzing and interpreting the PIP's results. Accurate PIP data analysis is critical because the MCO will implement changes in treatment and operations based on the results of a PIP.

Criteria

The review examines the appropriateness of, and the adherence to, the statistical analysis techniques defined in the data analysis plan. Interpretation and analysis of the study data should be based on continuous improvement philosophies and reflect an understanding that most problems result from failures of administrative or delivery system processes. Interpreting the data should involve developing hypotheses about the causes of less-than-optimal performance and collecting data to validate the hypotheses.

Potential Sources of Supporting Information

- Baseline project indicator measurements
- Repeat project indicator measurements
- Industry benchmarks
- Analytic reports of PIP results by the MCO

Assessment

Examine the calculated plan performance on the selected measures. To review the data analysis and results of the study, consider the following:

A. Is the analysis of the findings conducted in accordance with the data analysis plan?

B. Are numerical results and findings presented in an accurate, clear, and easily understood manner?

C. Does the analysis identify:

- Initial and repeat measurements of project outcomes?
- Realistic and unambiguous targets for the measures?
- The statistical significance of any differences between the initial and repeat measurements?
- Factors that influence the comparability of initial and repeat measurements?
- Factors that threaten the internal or external validity of the findings?

D. Does the analysis of the study data include an interpretation of the extent to which its PIP is successful and what follow-up activities are planned as a result?

Step 8: Assess the MCO's Improvement Strategies

In this step, the reviewer determines the appropriateness of the strategy for achieving true improvements. Real, sustained improvements result from a continuous cycle of measuring and analyzing performance, and developing and implementing system-wide improvements. Actual improvements depend on thorough analysis and implementation of appropriate solutions.

An improvement strategy is defined as an intervention designed to change behavior at an institutional, practitioner, or beneficiary level. The effectiveness of the intervention activity or activities is determined by measuring the MCO's change in performance, according to predefined quality measures.

Criteria

Interventions are key to a PIP's ability to bring about improved health care outcomes. Appropriate interventions must be identified and/or developed for each PIP to assure the likelihood of effecting measurable change.

If repeat measures indicate that quality improvement actions were not successful (i.e., did not achieve significant improvement), the problem-solving process should begin again with data analysis to identify possible causes and propose and implement solutions. If the quality improvement actions were successful, the new processes should be standardized and monitored.

Potential Sources of Supporting Information

- Current project baseline data
- Previous project data (if available)
- Results of clinical and literature research
- Project evaluation results completed by evaluators

Assessment

To assess the MCO's Improvement Strategies, consider the following questions:

A. Are the interventions related to causes/barriers identified through data analysis and quality improvement processes?

1. Interventions should be based on a root cause analysis of the problem the PIP addresses. It is expected that interventions associated with improvement on quality indicators will be system interventions (i.e., educational efforts, changes in policies, targeting of additional resources, or other organization-wide initiatives to improve performance). Interventions that might have some short-term effect, but that are unlikely to induce permanent change (such as a one-time reminder letter to physicians or beneficiaries) are insufficient.

2. An MCO is not required to demonstrate conclusively (e.g., through controlled studies) that a change in an indicator is the effect of its intervention; it is sufficient to show that an intervention occurred that might reasonably be expected to affect the results. Nor is the MCO required to undertake data analysis to correct for secular trends (e.g., changes that reflect continuing growth or decline in a measure because of external forces over an extended period). The MCO should be able to demonstrate that its data have been corrected for any major confounding variables with an obvious impact on the outcomes. The MCO's interventions should reasonably be determined to have resulted in measured improvement.

B. Are the interventions sufficient to be expected to improve processes or outcomes?

C. Are the interventions culturally and linguistically appropriate? For example, a mailing in English at 12th grade level to members of a predominately Chinese language population would not be appropriate. More information on culturally and linguistically appropriate services may be found at the following website: http://minorityhealth.hhs.gov/templates/browse.aspx?lvl=2&lvlID=15.

Step 9: Assess the Likelihood that Reported Improvement is "Real" Improvement

In this step, the reviewer determines the likelihood that the results of the PIP are accurate. It is important to determine if a reported change represents "real" change or is a result of a short-term event unrelated to the intervention, or simply random chance. Therefore, the EQRO must assess the probability that a reported improvement is a true improvement.

Criteria

"Real improvement" can be assessed in several ways, but is most confidently assessed by calculating the degree to which an intervention is statistically "significant". This protocol requires the EQRO to assess the extent to which any change in performance reported is statistically significant; however, it does not specify a specific level of statistical significance that must be met. States may establish a required level of statistical significance for findings to be accepted as valid. The EQRO should state in its final report which findings do not meet the required level of statistical significance.

Potential Sources of Supporting Information

- Baseline and repeat measures on quality indicators
- Tests of statistical significance calculated on baseline and repeat indicator measurements
- Benchmarks for quality specified by the State Medicaid agency or found in industry standards

Assessment

Review documents to determine the extent to which improvement occurred. Through repeated measurement of the quality indicators selected for the project, meaningful change in performance relative to the performance observed during baseline measurement must be demonstrated. The repeat measurement should use the same methodology as the baseline measurement, unless the baseline data was collected for the entire population at risk; the repeat measurement may then use a reliable sample. Performance using the identified indicators can be measured by collecting information on all individuals, encounters or episodes of care to which the indicator is applicable (a census) or by collecting information on a representative subset of individuals, encounters, providers of care, etc. Consider the following questions:

A. Are there any documented improvements in processes or outcomes of care?

B. Does the improvement in performance appear to be the result of the planned quality improvement intervention?

C. Is there any statistical evidence that any observed performance improvement is true improvement?

Step 10: Assess Sustainability of the Documented Improvement

Real change is the result of changes in the fundamental processes of health care delivery and is most valuable when it offers demonstrable sustained improvements. In contrast, a spurious "one-time" improvement can result from unplanned accidental occurrences or random chance. This step in the protocol is to determine if the real change is sustainable.

Criteria

Improvement must demonstrate repeated improvements or the likelihood of repeated improvements.

Potential Sources of Supporting Information

- Baseline and first repeated measurements on quality indicators
- Additional measurements on quality indicators made after the first repeat measurement

Assessment

Review of the re-measurement documentation is required to assure the improvement on a project is sustained. Consider the following question:

A. Is sustained improvement demonstrated through repeated measurements over comparable time periods?

Measurements of the outcomes are repeated after the first measurement following implementation of the intervention. Because of random year-to-year variation, population changes, and sampling error, performance on any given measure may decline in the second measurement. However, when all measurements for a given review are taken together, this decline should not be statistically significant.

ACTIVITY 2: VERIFY STUDY FINDINGS (OPTIONAL)

This activity is optional because verifying actual PIP study findings is a resource intensive activity that may not be feasible. If the PIP uses HEDIS® measures that have been certified by a third party, this step may not be needed. However, guidelines for conducting this optional activity are provided here.

Criteria

In addition to reviewing the methodology and findings of a PIP, States may request the EQRO verify the actual data produced to determine if the initial and repeated measurements of the quality indicators are accurate. This activity may not be feasible to perform for every (or even some) PIPs. Verification activities can provide added confidence in reported PIP results as they provide greater evidence that the findings are accurate and reliable. Therefore, this activity is included in this protocol as an optional activity that a State may elect to have the EQRO conduct on an ad hoc basis when the State has special concerns about data integrity.

Potential Data Sources Needed for Verification Activities:

- Current project data and findings
- Depending upon the source of the PIP data:
  - MCO administrative data
  - Beneficiary interviews and surveys
  - An assessment of the MCO's Information System (see Appendix V)

Assessment

The key focus in this activity is validating the processes through which data needed to produce quality measures were obtained, converted to information, and analyzed. Assessing the algorithm together with the integrity of the MCO's information system and encounter data will provide a strong indication of the accuracy of the MCO's reported quality measures. The algorithm for converting the information and analyzing it is verified in Activity 2, Step 6 of this protocol. The methods used to verify how the data were collected depends on whether the data are obtained through review and abstraction of medical records or produced through an automated information system.

Verification of quality measures produced through medical record review can be achieved by conducting a re-abstraction of a small subset (validation sample) of the reviewed records. Data retrieval and analysis should be conducted on a small scale, with the validation sample following the same abstracting rules as the original study. Statistical correlations will be made between the validation sample and the original study data. A wide variety of statistical methods can be applied to assess the degree of correlation between the study and validation measures. Two recommended methods are the Pearson correlation coefficient for continuous data (e.g., age, income, etc) and the Kappa statistic for categorical data (e.g., gender, race, etc.).

Verification of data obtained though MCO-automated information system is a reflection of three phenomena:

1. Soundness of the algorithm used to produce quality measures from its information system;
2. Integrity (completeness and accuracy) of the MCO's information system at capturing enrollee information; and
3. Accuracy of the information translated from source documents (e.g., an enrollee's medical record) into automated data in the MCO's information system.

These three activities can be performed by one or more of the following methods:

- Review the assessment of the MCO's information system and any validations of MCO encounter data that the State has produced as described in Appendix V.
- Review the results of another Protocol or EQR activity (e.g., validating encounter data, validating performance measures, or assessing an MCO's compliance with standards for MCO information system specified by the State Medicaid agency or other organization such as a private accrediting organization.
- Review the MCO's own recently completed assessment of the MCO's information system and validation of its encounter data from the MCO, the State Medicaid agency, or other organization identified by the MCO.
- In the event that no current evaluation of an MCO's information system or encounter data exists, the State may choose to contract this important function to fulfill this requirement to validate its MCO PIPs.

ACTIVITY 3: EVALUATE AND REPORT OVERALL VALIDITY AND RELIABILITY OF PIP RESULTS

Following the completion of Activity 1 and Activity 2, the EQRO will assess the validity and reliability of all findings to determine whether or not the State has confidence in the MCO's reported PIP findings. As studies always have weaknesses, the EQRO will need to accept threats to the accuracy of the PIP, and determine PIP generalizability as a routine fact of QI activities.

The EQRO will report findings to the State. The PIP validity report should include the description of the PIPs that were validated and the findings of the EQRO's validation review. Because determining threats to validity, reliability, and PIP design is sometimes a judgment call, the EQRO can report a level of confidence in its findings. Examples of levels that can be reported to the State include:

- High confidence in reported MCO/ PIP results;
- Confidence in reported MCO/ PIP results;
- Low confidence in reported MCO PIP results; or
- Reported MCO PIP results not credible.

The EQRO and the State must include the actual results of the PIPs in the final EQRO technical report for submission to CMS.

REFERENCES

Quality Improvement System for Managed Care (QISMC)

Health Care Quality Improvement Studies in Managed Care Settings: A Guide for State Medicaid Agencies (National Committee for Quality Assurance (NCQA))

A Health Care Quality Improvement System for the Medicaid Managed Care, A Guide for States (Health Care Financing Administration (HCFA))

Framework for Improving Performance, From Principles to Practice (Joint Commission on Accreditation of Healthcare Organizations (JCAHO)

1990-2000 Standards for Health Care Networks (SHCN) (JCAHO)

NCQA 1997, 1998, and 1999 Standards for Accreditation of Managed Care Organizations and NCQA 1999 Standards for Accreditation of Managed Behavioral Healthcare Organizations (MBHO)

Peer Review Organizations (PRO) 4th and 5th Scope of Work (SOW) (CMS)

*Please see EQR Protocols Appendix I for information about how these references were used to develop PIP protocols.

END OF PROTOCOL