Lessons from the Field: Identifying Measurement Challenges and Successes

Prepared for the Agency for Healthcare Research and Quality by L&M Policy Research, LLC with guidance from the Pediatric Quality Measure Program (PQMP) Grantees

Table of Contents

List of Acronyms	
·	
Lessons from the Field: Identifying Measurement Challenges and Successes	3
Introduction	
Challenges to Measurement	
Grantee Approaches	
Grantee Key Findings and Implications	8
References	13

List of Acronyms

AHRQ Agency for Healthcare Research and Quality

CheQ University of Florida/Child Health Quality Partnership

KI Key Informant

KII Key Information Interviews

NCINQ II National Committee for Quality Assurance/National Collaborative for Innovation in

Quality Measurement: Implementing and Improving

PQMP Pediatric Quality Measure Program

PQMP-LC Pediatric Quality Measure Program Learning Collaborative

QI Quality Improvement

Q-METRIC University of Michigan/Quality Measurement, Evaluation, Testing, Review, and

Implementation Consortium

RF Research Foci

Lessons from the Field: Identifying Measurement Challenges and Successes

Introduction

This lesson from the field report examines one Research Foci (RF) central to the Pediatric Quality Measures Program (PQMP) grantees' work. This RF broadly focuses on the measurement challenges and successes experienced at each level (state, health plan, practice, provider and patient), as depicted in Figure 1.

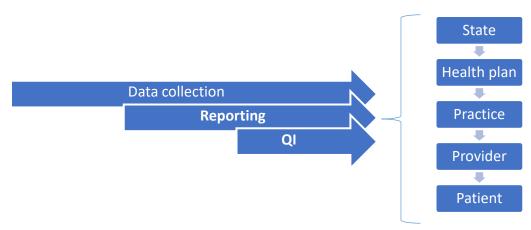


Figure 1: Measurement Challenges and Successes

The specific question is:

 What measurement (e.g., data collection, reporting, QI) challenges and successes are identified at different levels (e.g., state, health plan, practice, provider, patient levels)?

In examining this question across two distinct data sources—literature reviews and key informant interviews (conducted by the PQMP Learning Collaborative)—a set of key considerations emerged that relate to measurement challenges and success at different levels. Each of these key considerations and supporting findings from the literature and key informants are discussed below.

Availability of accurate, timely, and complete data.

The availability of accurate, timely and complete data remains a key challenge to quality measurement and reporting at multiple levels. This is an issue particularly when using measures at different levels as data sources differ across levels, with health plans generally using claims data for measure construction and hospitals or other providers more often relying on medical records data. Some measures may also require data elements from more than one source. Linking data across diverse sources (e.g., medical records, vital statistics, claims) to create a complete patient record is a solution that supports sound

quality measurement; it is also difficult and may not be feasible due to infrastructure, lags, confidentiality, policy, or other limitations, creating further challenges for measurement and reporting.

According to the RAND Corporation's evaluation of the uses of NQF-endorsed measures, "[t]he single most important factor cited as either facilitating or impeding the use of measures was the availability of data to construct performance measures" (Damberg et al., 2011, p.xv). While the majority of data used to measure quality come from health care claims submitted by providers, administrative claims are not intended to describe the non-billable care provided, to record and manage patient care, or capture how patients are affected by care received (Anthem Public Policy Institute, 2016). "Hybrid" data collection may sometimes be required, which refers to the combination of administrative data with information obtained from medical records (or patient experience surveys) and may generate more robust quality measures by increasing the number of component data elements (Berenson et al., 2013).

Key informants described successes and challenges in linking data from multiple sources for measurement and reporting. One key informant from a health plan noted successes in linking claims and medical records for measurement purposes but described challenges experienced in trying to link data on immunizations from state registries to claims data. A state medical director noted her state's success in linking claims and vital statistics records, allowing construction of measures for mothers and infants.

Complexity of specifications within and across entities.

Complexity, lack of standardization, or inconsistency in specifications within and across entities and levels, as well as over time, is a factor impeding use of and reporting on quality measures. In order to calculate measures such that the results are comparable, entities need to have clear measure definitions, specifications and appropriate data or a clear understanding of when alternate data sources can be used.

The authors of the Anthem Public Policy Institute 2016 issue brief on quality measurement in Medicaid managed care wrote that some measures are simply not "amenable" to computer programming (Anthem Public Policy Institute, 2016). Technical factors such as clarity and complexity of measure specifications and software capabilities complicate measurement activities (Ireys et al., 2015). Due to the complexities of the specifications, state grantees in the CHIPRA Quality Demonstration Grant program needed to contact CMS's technical assistance mailbox for queries on measure numerators, denominators, sampling methods, use of alternative data sources, coding systems, and continuous enrollment criteria (Christenson, 2017, p.190)

Based on challenges and complexity around the transition from ICD-9 to ICD-10 billing codes and using them for reporting, Caskey et al. (2014) found that 26 percent of all pediatric ICD-9-CM codes and 21 percent of pediatric patient visits were associated with "convoluted" ICD-10-CM codes. While the authors contend that the transition to ICD-10-CM should lead to an increased level of clinical detail included in each code, they wrote that the "convoluted" codes had the immediate potential to cause documentation inaccuracies, which could lead to financial loss from erroneous administrative data and billing errors as well as errors in surveillance. These changes could also

One of the key informants stressed the importance of clear and consistent measure specifications, noting that sometimes changes do not seem significant " ... until you actually see how much of a change that little tweak on the specifications actually had on the measure results."

lead to challenges in measuring performance over time, when measures rely on diagnostic codes that have changed.

Measure alignment across entities requiring reporting.

Alignment of measures across the different entities that require reporting (e.g., Medicaid, commercial payers, etc.) can improve measure uptake as well as performance reporting. Strategies that seek to streamline program requirements, align improvement objectives, and reduce operational/financial burden will enhance stakeholder buy-in and engagement in measurement activities.

A number of experts (Conway et al., 2013, Easter & Venkatesh, 2017, McGinnis et al., 2015) argue that alignment of measures encourages more coordinated quality improvement activities, reinforces the same set of incentives, and yields more consistent performance measurement. Additionally, Damberg et al. (2011) observed that "strong data infrastructure; provider trust in the measurement process and the evidence base of the measures; alignment of measures among reporting initiatives to minimize

reporting burden; relevance to members, consumers, and providers; and provider training on how to extract the data" were key parameters by which quality measurement can be successfully undertaken (p.xv).

Damberg et al. (2011) refer to "measure fatigue" in which the preponderance of measures and lack of alignment across national, state, and regional programs impedes the use of measures. In their issue brief on quality measurement in Medicaid managed care, the Anthem Public Policy Institute reported evidence of further idiosyncrasy in state and regional quality measure sets, where analysts found over 500 measures in use in 48 public programs, where only 20 percent of measures were used in more than one measure set, and many were used in a single state.

One of the key informants noted how her state was interested in alignment of measures across payers. She acknowledged that "there'll be instances in which we have measures for the Medicaid population that may not transcend across our commercial enrollees," but emphasized that their goal was alignment and ways to minimize data collection and reporting burden.

Importance of understanding variation in care and data.

Understanding variation in care processes and harmonizing related data collection is important for accurate measurement for quality improvement, regardless of the measurement level (state, health

plan, or provider). Variation in care processes can lead to variation in the way in which data elements are operationalized or recorded. To ensure measures are comparable, data collection must be clearly delineated so that measures represent the same underlying concepts.

Berenson et al. (2013) described site-to-site variation in defining the medical term "shock" and how it may lead to biases in assessing performance in sepsis quality measures (Berenson et al., 2013). The authors noted

One key informant described challenges faced by a managed care plan when developing a registry to identify hypertensive individuals. Blood pressure was not being clinically measured in a consistent manner which compromised data quality. To address this challenge, the plan worked with providers in one region to determine standardized procedures to encourage best practices for taking patients' blood pressure to support consistent data capture for hypertensive individuals.

that this variability in the quality of data in medical records confounds efforts to appropriately profile and benchmark performance across sites.

The remainder of this *Lessons from the Field* provides examples of how the work of three grantees specifically relates to the considerations described above for one or more of their pediatric quality measures. For each key consideration, the grantees described: (1) the challenges they faced during implementation, (2) the approach(es) they took to address the challenges, and (3) their team's specific findings and implications for measure implementation.

Challenges to Measurement

The grantees identified a number of challenges related to measurement. While not all grantees faced the same challenges, there were a few commonalities across projects related to the key considerations. Two of the grantees' projects lacked access to the clinical data needed to support calculating a measure or to further understand patient characteristics in support of their QI efforts at the health plan or provider level. Despite measure alignment across reporting, several grantees struggled to engage providers in QI efforts. Grantees reported that providers either prioritized projects that impacted larger patient populations or participants generally lacked the resources to engage in QI efforts. Several examples are presented in Figure 2.

Figure 2. Examples of Grantee Measurement Challenges, by Key Consideration

Availability of accurate, timely, and complete data

- NCINQ: obtaining/sharing clinical data across levels to accurately construct measures.
- Q-METRIC: measure uses claims data, and is reported and assessed only annually (at state and health plan level), which is not frequent enough to support QI efforts.
- •CHeQ: availability of data from multiple sources, at the provider and health plan level, to understand the clinical profiles of children assessed by the measure.

Complexity of specifications within and across entities

- NCINQ: identified issues with how health plans interpreted measure specifications, leading to results that did not meet validation criteria.
- Q-METRIC: struggled with re-specifying and validating measures (previously endorsed by NQF using ICD-9 codes) using ICD-10 codes to support accurate reporting over time.

Measure alignment across entities requiring reporting

- NCINQ: despite benefits of using nationally reported measures at the plan level, encountered low engagement from providers due to smaller numbers of affected patients.
- Q-METRIC: achieved measure alignment across all levels but faced challenges engaging providers in QI efforts.

Importance of understanding variation in care and data

 NCINQ: variation in how depression screening and monitoring was being performed and documented at the practice level as well as variation in provider workflows to address follow-up care and monitoring.

Grantee Approaches

The grantees' approaches to addressing the key measurement challenges relied both on quantitative data analyses, as well as stakeholder engagement activities. Selected examples are shown in Figure 3.

Figure 3. Grantees used varied approaches to addressing measurement challenges

Data analysis

- •The **CHeQ** team conducted subgroup analyses of children taking antipsychotics to better understand variations by geography, race, ethnicity and sex to help them target and tailor interventions at the provider and health plan levels.
- In order to implement their appropriate antibiotic prophylaxis among children with sickle cell anemia measure at multiple levels, Q-METRIC developed an algorithm that linked state Medicaid claims to health systems.
- •The Q-METRIC team found that the public or commercially available code conversion tools created inaccurate results, so they went through a comprehensive manual review to translate the appropriate antibiotic prophylaxis among children with sickle cell anemia measure from ICD-9 to ICD-10 codes in order to accurately reflect the measure specifications and consistently measure performance across entities and over time.

Stakeholder interactions

- •The **NCINQ** team convened a learning collaborative focused on adolescent depression with five Medicaid health plans across the country. The collaborative helped participating plans develop approaches to aggregate disparate data sources needed to understand depression care provided and to report the measures. The participating health plans also worked with individual provider practice sites to collect data and conduct quality improvement projects focused on depression care for adolescents.
- Q-METRIC is working closely with the state of Michigan to help support the development and implementation of a novel QI program focused on providing incentives to health plans for improved performance on their sickle cell antibiotic prophylaxis measure.
- •The **CHeQ** team led a broad stakeholder collaborative effort to develop the infrastructure for linked electronic health record and Medicaid claims data in Florida to support QI efforts for metabolic monitoring for children taking antipsychotics.

Grantee Key Findings and Implications

Based on their analytic and stakeholder activities, grantees produced findings for each of the key considerations about measurement challenges and successes. These findings expand the evidence base and strengthen the connection between measurement and improvement. Across levels, several grantees found that accurate measurement required using data from multiple sources, including use of data from medical records to increase timeliness of measures and enhance measurement. Grantees also recognized that careful specification of measures was critical to consistent measurement across entities and levels. More of the key findings and the implications for measurement and improvement efforts are presented in the following series of tables.

Figure 4: The availability of accurate, timely and complete data remains a key challenge to quality measurement and reporting at multiple levels, particularly when measures require data elements not available in claims data.

Key Findings

- The NCINQ team's depression measures require documentation
 of depression screening and follow-up symptom monitoring
 assessments. These data reside in the provider EHR systems.
 Plans' lack of access to providers' EHRs made it difficult to capture
 screening results necessary to calculate the measure.
 - Health plans in the NCINQ collaborative were able to use a variety of strategies to aggregate different data sources to get a more complete picture of quality of adolescent depression care.
- Leveraging the infrastructure developed by Florida's Medicaid program and partner health systems, the CHeQ team was able to outline and implement a process to address issues of data linkage, data privacy and access to more real-time data. This allowed for the development of a metabolic profile of children taking antipsychotic medications.
 - Based on the Florida experience, the team is pilot testing a partnership in Texas to develop a similar, though smaller-scale, infrastructure.
- While there is a lag in claims data availability at the state level, the Q-METRIC team found very high reliability in the performance scores calculated for plans using state data and scores calculated by plans using their own data. The latter were available in a timelier manner to support QI efforts.
 - At the health system level, the team found that data were available from medical records to indicate whether a prescription was written, but not whether the prescription was filled (only available in claims).

- ⇒ Augmenting claims data with clinical data may be needed to obtain required data elements and accurately measure patient outcomes.
 - Health plans can use a variety of strategies, including data sharing agreements, partnerships with data aggregators and EHR vendors, leveraging of state/regional HIEs as well as educating providers about the utility of using standard codes to capture clinical services at the point of care.
 - Health plans also suggested that their screening focused
 QI efforts might benefit from an additional resource
 tool (such as a value set of existing billing codes); a value
 set could be used to help them understand baseline
 results by provider and population subgroup and help
 them develop appropriate incentives.
- ⇒ While time and resources are required, it is possible to develop statewide or regional infrastructure that links data sources to support measurement and reporting of health outcomes. States play an important role in facilitating the development of systems to enable and promote data sharing.
- ⇒ To overcome lags in claims data availability, states may be able to partner with plans to use plan-level claims data for calculating more timely performance scores.
 - Plans may also be able to use monthly calculations on a rolling basis to monitor QI efforts.

Figure 5: Complexity, lack of standardization, or inconsistency in specifications within and across entities and levels is a factor impeding use of and reporting on quality measures.

Key Findings

- Although health plans in the learning collaborative submitted data using a data collection tool, the NCINQ team found that some plans were interpreting and applying the HEDIS specifications differently, leading to results that did not meet the validation criteria.
 - The team's provision of additional guidance on measure definitions and data capture reduced the need for interpretation and increased comparability of results.
- The NCINQ team found differences in the exclusion criteria related to diagnoses between the HEDIS plan-level depression screening measure and the CMS provider-level depression screening measure.
 - Feedback from stakeholder panels was used to inform standardization of the measures across reporting entities and levels.
- After re-specifying their appropriate antibiotic prophylaxis among children with sickle cell anemia measure using ICD-10 codes, the Q-METRIC team shared the code with partner health plans and two state Medicaid programs.
 - The health plans and states calculated performance scores, testing the revised specification to ensure that it could be used to measure performance across entities and over time.

- ⇒ Even when using measures for QI purposes, it is important to have detailed specifications, clear guidelines for how data should be reported, and an audit/validation process to ensure adherence to specifications and measure comparability across entities and levels.
- Standardization of measure definitions across reporting requirements can improve consistency of measure collection and reporting across levels.
- ⇒ In scenarios that require changes or updates to measure specifications such as the transition from ICD-9 to ICD-10 codes
 it is important to thoroughly test measures to assess performance across levels.

Figure 6: Alignment of measures across the different entities that require reporting (e.g., Medicaid, commercial payers, etc.) can improve measure uptake as well as performance reporting.

Key Findings

- The NCINQ team found that state requirements to use national program specifications for their three depression measures ensured alignment of data elements and coding. Having a financial incentive for reporting the metabolic monitoring measure further contributed to high uptake and performance results that were widely available.
 - This availability allowed the team to examine and compare performance rates for plans participating in the collaborative, other plans in the state, and plans nationally.

- ⇒ State reporting requirements and incentives help to focus attention on performance measurement and reporting. This may be particularly important in creating the impetus for plans to focus resources for measures related to small but vulnerable populations.
- ⇒ Increasing measure uptake for providers may require multiple strategies including individual provider feedback reports and comparisons with other providers as well as provider education on the importance of following guidelines concerning recommended services for the patient population.

Figure 7: Understanding variation in care processes and harmonizing related data collection is important for accurate measurement for quality improvement.

Key Findings

- As part of the NCINQ collaborative, plans found that there was substantial variation in how practices performed and documented depression screening and monitoring. This underlying variation could impact use of the measure across levels by affecting plans' ability to assess provider performance on the measure.
 - For example, in some cases, plans were not receiving clinical data from practices, so performance results inappropriately indicated no screening or assessments were occurring.
 - Even when providers were documenting results of depression assessments in their EHRs, often the results were not captured in structured fields using standard coding specifications.

- ⇒ To measure patient outcomes, there is a need for more consistent documentation in structured fields within data systems that can be queried and shared across different care settings for quality measurement and improvement purposes.
- Plans' education of providers may help improve the workflow for screening and assessment measures (e.g., depression), best practices for conducting screenings, and documentation over time. These changes may contribute to better measurement of patient outcomes.
- ⇒ Development of digital quality measures can help ensure that measures are consistently calculated. Digital measures include use of standard terminologies and specific codes with the logic of the measure calculation provided in both a narrative human-readable form and executable files that computers can use to 'read' the logic.

References

- Agency for Health Care Research and Quality.

 https://www.ahrq.gov/sites/default/files/wysiwyg/policymakers/chipra/factsheets/chipra_1415
 -p011-1-ef.pdf accessed November 4, 2018.
- Anthem Public Policy Institute. (2016). The "Nuts and Bolts" Behind Quality Measurement in Medicaid Managed Care. Retrieved from https://www.anthempublicpolicyinstitute.com/blog/the-nuts-and-bolts-behind-quality-measurement-in-medicaid-managed-care-1
- Berenson, R. A., Pronovost, P. J., Krumholz, H. M. (2013). Achieving the Potential of Health Care

 Performance Measures. Retrieved from http://webarchive.urban.org/publications/412823.html
- Bian J, Loiacono A, Sura A,..Shenkman E, Hogan W. Implementing a hash-based privacy-preserving record linkage tool in the OneFlorida clinical research network. JAMIA Open. 2019;2(4):562–569. Published 2019 Sep 27. doi:10.1093/jamiaopen/ooz050
- Caskey, R., Zaman, J., Nam, H., Chae, S. R., Williams, L., Mathew, G., ... Boyd, A. D. (2014). The Transition to ICD-10-CM: Challenges for Pediatric Practice. *Pediatrics*, *134*(1), 31–36. doi:10.1542/peds.2013-4147
- Christensen A. L., Petersen, D. M., Burton R. A., Forsberg V. C., & Devers K. J. (2017). What Factors Influence States' Capacity to Report Children's Health Care Quality Measures? A Multiple-Case Study. *Maternal and child health journal*, 21(1), 187-198. https://doi.org/10.1007/s10995-016-2108-8
- Conway, P. H., Mostashari, F., & Clancy, C. (2013). The Future of Quality Measurement for Improvement and Accountability. *JAMA*, *309*(21), 2215-2216. doi:10.1001/jama.2013.4929
- Damberg, C.L., Sorbero, M.E., Lovejoy, S.L., Lauderdale, K.E., Wertheimer, S.R., Smith, A.K., Waxman, D., & Schnyer, C. (2011). An Evaluation of the Use of Performance Measures in Health Care.

 Retrieved from https://www.rand.org/pubs/technical_reports/TR1148.html
- Easter, B., & Venkatesh A., (2017). Measuring and Improving Quality of Care. In J. L. Wiler, J. M. Pines & M. J. Ward (Eds.), *Value and Quality Innovations in Acute and Emergency Care* (pp.18-26). Cambridge: Cambridge University Press.
- Ireys, H., Zickafoose, J., Petersen, D., Christensen, A. (2015) National Evaluation of the CHIPRA Quality Demonstration Grant Program: Final Project Report: Rockville, MD, 2015 (AHRQ Publication No. 16-0007-EF). Retrieved from https://www.ahrq.gov/policymakers/chipra/demoeval/what-we-learned/final-report/index.html

- Maayan L, Correll CU: Weight gain and metabolic risks associated with antipsychotic medications in children and adolescents. Journal of Child and Adolescent Psychopharmacology 21:517–535, 2011
- McGinnis, J. M., Malphrus, E., & Blumenthal, D. (Eds.). (2015). Vital signs: Core Metrics for Health and Health Care Progress. Retrieved from http://nationalacademies.org/HMD/reports/2015/vital-signs-core-metrics.aspx
- Shenkman E, Hurt M, Hogan W, et al. OneFlorida Clinical Research Consortium: Linking a Clinical and Translational Science Institute With a Community-Based Distributive Medical Education Model. *Acad Med.* 2018;93(3):451–455. doi:10.1097/ACM.000000000002029
- Yuan J, Malin B, Modave F, ...Shenkman E, Bian J. Towards a privacy preserving cohort discovery framework for clinical research networks. *J Biomed Inform*. 2017;66:42–51. doi:10.1016/j.jbi.2016.12.008