Variation in Outcomes of Quality Measurement by Data Source

abstract

OBJECTIVE: To evaluate selected Children’s Health Insurance Program Reauthorization Act claims-based quality measures using claims data alone, electronic health record (EHR) data alone, and both data sources combined.

METHODS: Our population included pediatric patients from 46 clinics in the OCHIN network of community health centers, who were continuously enrolled in Oregon’s public health insurance program during 2010. Within this population, we calculated selected pediatric care quality measures according to the Children’s Health Insurance Program Reauthorization Act technical specifications within administrative claims. We then calculated these measures in the same cohort, by using EHR data, by using the technical specifications plus clinical data previously shown to enhance capture of a given measure. We used the κ statistic to determine agreement in measurement when using claims versus EHR data. Finally, we measured quality of care delivered to the study population, when using a combined dataset of linked, patient-level administrative claims and EHR data.

RESULTS: When using administrative claims data, 1.0% of children (aged 3–17) had a BMI percentile recorded, compared with 71.9% based on the EHR data (κ agreement [k] = 0.01), and 72.0% in the combined dataset. Among children turning 2 in 2010, 20.2% received all recommended immunizations according to the administrative claims data, 17.2% according to the EHR data (k = 0.82), and 21.4% according to the combined dataset.

CONCLUSIONS: Children’s care quality measures may not be accurate when assessed using only administrative claims. Adding EHR data to administrative claims data may yield more complete measurement.

Pediatrics 2014;133:e1676–e1682
The Children’s Health Insurance Program Reauthorization Act of 2009 (CHIPRA), intended to increase the number of children with health insurance, also mandated that a set of pediatric care quality measures be selected.\textsuperscript{1,2} An initial set of 24 CHIPRA core measures, chosen for their validity, importance to health outcomes, and ability to be evaluated by using states’ Medicaid administrative claims data, were identified by experts to meet this mandate.\textsuperscript{1,3–7}

States are encouraged to report on the CHIPRA measures to assess the quality of care provided to children in their Medicaid program and Children’s Health Insurance Program (CHIP).\textsuperscript{4} Methods for operationalizing each measure are given in the CHIPRA Technical Specifications and Resource Manual.\textsuperscript{8} In most cases, these methods contain parameters for defining the population of children to include in the rate denominator, and a set of diagnostic and procedure codes from administrative claims data to identify provided care (the numerator). A “hybrid” method using manual chart review on a sample of children is also given for defining some of the measures; however, such reviews are rarely feasible for states to conduct, as obtaining medical record data can be difficult, costly, and time-consuming.\textsuperscript{8,9}

The CHIPRA measures are an important step toward standardizing pediatric care quality assessments. However, limitations to using administrative claims data for such purposes have been shown in previous research assessing care quality.\textsuperscript{10–13} Although adding data from discrete electronic health record (EHR) data fields shows promise in improving the accuracy of claims-based quality measurement,\textsuperscript{11,14,15} assessments using EHR data alone may have limitations.\textsuperscript{16–19} For example, in a study of the quality of care provided to adults with diabetes, neither Medicaid administrative claims nor EHR data alone yielded a complete capture of provided services: 12% of cholesterol screenings received were documented in Medicaid claims but not the EHR, 49% were documented in the EHR but not in Medicaid claims, and only 39% were documented in both data sources.\textsuperscript{10}

After the CHIPRA measures were initially developed, Gold et al\textsuperscript{12} determined which of these claims-based measures could feasibly be calculated by using electronically abstracted EHR data, and what modifications to the claims-based technical specifications were needed, if any, to enable such EHR-based measurement. Building directly on that landmark study, the present analysis sought to evaluate the accuracy of selected CHIPRA claims-based measures, when based on administrative claims data alone, EHR data alone, or both data sources combined. We also built on this previous work by assessing the quality of care delivered to a single cohort of continuously insured children who were established patients within the OCHIN network of community health centers.

**METHODS**

**Data Sources**

**Oregon Health Plan Administrative Data**

Oregon’s Medicaid and CHIP insurance programs are both operated by the Oregon Health Plan (OHP). The OHP pays clinics for provided health care, and thus receives claims (ie, diagnosis and procedure codes) for all OHP-insured children in Oregon.

**EHR Data**

“OCHIN, Inc.” originally the Oregon Community Health Information Network, provides and maintains a comprehensive Epic EHR infrastructure for member clinics. All sites share a single, linked EHR and patients have 1 unique record; data are stored in a central repository that can be searched electronically.\textsuperscript{19} All of the EHR data used in this study were pulled by using automated extraction from 46 OCHIN clinics in Oregon that had the full EHR implemented before July 1, 2009 (6 months before the study year).

**Study Population**

We selected a study cohort of children and adolescents (aged $\leqslant$21 years) who were continuously insured by OHP and had at least 1 outpatient visit at a study clinic in 2010 ($n = 22,925$ children). We deterministically linked records for these children who appeared in both the OHP claims and the OCHIN EHR datasets by using the child’s unique OHP client identification number. For all of the measures calculated here, we used the CHIPRA technical specifications’ definition of continuously insured: briefly, a child or adolescent must have 12 months of coverage with no more than one 45-day gap. Depending on the measure, different time periods apply for when the child/adolescent had to have continuous insurance; for example, the BMI percentile measure requires continuous insurance during the measurement year, whereas the childhood immunization measure requires continuous insurance during the year before the child’s second birthday.\textsuperscript{8}

**Measures**

We selected a subset of the CHIPRA measures for analysis based on those previously identified as feasible for assessment using EHR data and those most applicable to the children in our study population, including BMI percentile, recommended immunizations for children and adolescents, well-child visits, and *Chlamydia* screening for sexually active women.\textsuperscript{12} We did not include 2 current and feasible measures due to sample size: developmental
screening had a small numerator \((n \sim 221)\) and follow-up care for children prescribed attention-deficit/hyperactivity disorder medication had a small denominator \((n \sim 151)\). The prenatal measures were beyond the scope of this project.

In analyses within administrative claims data, we used the exact measure specifications from the Initial Core Set of Children’s Health Care Quality Measures: Technical Specifications and Resource Manual for Federal Fiscal Year 2011.\(^8\) Optional exclusions for the measures were not used. In the assessments using electronically pulled EHR data, we calculated the measures by using the technical specifications (ie, CPT codes, and so forth) plus relevant clinical data (ie, historical immunizations, primary reason for the visit, and so forth) as outlined by Gold et al.\(^12\) See Table 1.

### Analysis

First, we described the sociodemographic characteristics of the study population. Second, we assessed rates of receipt for each of the selected quality measures among the measure-specific subgroup of children in the study population, by using the administrative claims data, and again by using the EHR data (which included the additional fields unique to the EHR dataset described previously). Third, we tested for agreement in quality measurement using the administrative claims dataset versus the EHR dataset, by using the \(\kappa\) statistic. Fourth, we assessed care quality by using a linked dataset that combined the administrative claims and EHR data. Finally, among all services documented in the combined dataset, we determined the percentage found in claims alone, the EHR alone, or in both data sources. All analyses were conducted by using SAS software, version 9.3 (SAS Institute, Inc, Cary, NC). This study was approved by the institutional review board at Oregon Health & Science University.

### RESULTS

#### Demographics

The mean age of the children in our study population was 8.5 years. More than half were white (52.4%) and nearly half were Hispanic (47.3%). The main languages spoken were English (45.5%) and Spanish (41.0%) (Table 2).

### Measures Assessed by Using Claims Data Alone, EHR Data Alone, or a Combined Claims/EHR Dataset

#### BMI

In the administrative claims data, 1.0% of children (aged 3–17) in our study population had a BMI percentile recorded in

<table>
<thead>
<tr>
<th>Table 1 Specifications Used to Assess Selected CHIPRA Measures by Using Administrative Claims and EHR Data</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>CHIPRA Measure</strong></td>
</tr>
<tr>
<td>BMI percentile documentation</td>
</tr>
<tr>
<td>Immunization status: child</td>
</tr>
<tr>
<td>Immunization status: adolescents</td>
</tr>
<tr>
<td>Well-child care: infants</td>
</tr>
<tr>
<td>Well-child care: children</td>
</tr>
<tr>
<td>Well-child care: adolescents</td>
</tr>
<tr>
<td>Annual Chlamydia screening</td>
</tr>
</tbody>
</table>

\(^*\) All measures calculated from the same cohort of children: those continually insured by the OHP, which administers both Oregon’s Medicaid Program and CHIP, with \(\geq 1\) outpatient visit in 2010 at 1 of the 46 Oregon clinics that had fully implemented OCHIN’s EHR system before July 9, 2009.

\(^8\) Data to meet the technical specifications were electronically pulled from the OCHIN EHR.
Among children who turned 2 in 2010, 20.2% had received all their recommended immunizations, when measured by administrative claims data; 17.2% met the measure according to the EHR data (k=0.82). In the combined dataset, 21.4% had received all recommended immunizations. Among adolescents who turned 13 in 2010, 51.8% had received both a tetanus, diphtheria, acellular pertussis/tetanus, diphtheria and meningococcal immunization according to the administrative claims data, and 48.4% when assessed in the EHR data (k = 0.82). In the combined dataset, 54.6% of adolescents were up-to-date.

Well-Child Visits

Among children who turned 15 months of age in 2010, 59.5% had received at least 6 well-child visits when measured by OHP administrative claims data. When measured in the EHR data, 60.5% were up-to-date on this measure (k = 0.49). In the combined dataset, 72.2% of children were up-to-date. The same pattern held true for well-child visits for 3- to 6-year-olds and for adolescents: similar results were found in each data source when assessed separately, and the rate was higher when the data sources were combined.

Chlamydia screening

In the assessments using OHP administrative claims data, 41.1% of sexually active adolescent girls and young women aged 16 to 20 had a Chlamydia test; 76.0% of eligible adolescent girls and young women had a Chlamydia test according to the OCHIN EHR data (k = 0.26). In the combined dataset, 78.6% of the eligible population had been screened (Table 3).

DISCUSSION

Increasingly, care quality assessed by using standardized methods, such as the CHIPRA core measures, will be used to calculate incentives or reimbursements. For example, in Oregon, coordinated care organizations will receive incentive payments based on the quality of care provided to Medicaid patients enrolled in their programs.20 Thus, it is increasingly important to understand the limitations inherent to measuring care quality by using different data sources, to ensure that these data provide an accurate reflection of the care provided.

This study found significant differences between claims and EHR data sources depending on the measure assessed: BMI was recorded in the EHR far more often than in claims, and immunizations were often recorded in claims data but not found in the EHR data. There are logical explanations for the lack of agreement between these 2 data sources. For example, the lack of information in claims about BMI assessment highlights a limitation of claims data: these data represent billed services only. Conversely, immunizations may be received at multiple care sites and billed to patients’ insurance plans, but not reported back to or reflected in their primary care medical home’s EHR. Our findings suggest that most current methods for assessing care quality are limited because they are usually based on information obtained from 1 data source. If quality measurements use only administrative claims data, which document billed services and associated diagnoses, they are not capturing all of the care received or services provided.21–24 Similarly, quality measurements that use EHR data alone may not obtain a complete picture of all care patients receive, as EHR data often only include data from 1 clinic system, and patients may seek care at different locations. In this study population, combining administrative claims and EHR data yielded the most complete capture of data on the quality of care provided to children. These findings illustrate that different

### TABLE 2

<table>
<thead>
<tr>
<th>Child Characteristics</th>
<th>Children Continuously Enrolled in OHP With ≥1 Outpatient Visitb (n = 22,925)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Age in years, mean (SD)</td>
<td>8.5 (6.0)</td>
</tr>
<tr>
<td>Race</td>
<td></td>
</tr>
<tr>
<td>Black</td>
<td>1775 (7.7)</td>
</tr>
<tr>
<td>Asian/Pacific Islander/Native Hawaiian</td>
<td>1220 (5.3)</td>
</tr>
<tr>
<td>American Indian/Alaskan Native</td>
<td>182 (0.8)</td>
</tr>
<tr>
<td>White</td>
<td>12,023 (52.4)</td>
</tr>
<tr>
<td>Unknown/Refused</td>
<td>7,726 (33.7)</td>
</tr>
<tr>
<td>Ethnicity</td>
<td></td>
</tr>
<tr>
<td>Hispanic</td>
<td>10,834 (47.3)</td>
</tr>
<tr>
<td>Non-Hispanic</td>
<td>5,442 (23.7)</td>
</tr>
<tr>
<td>Not Collected/Unknown</td>
<td>6,649 (29.0)</td>
</tr>
<tr>
<td>Language</td>
<td></td>
</tr>
<tr>
<td>English</td>
<td>10,426 (45.5)</td>
</tr>
<tr>
<td>Spanish</td>
<td>9,397 (41.0)</td>
</tr>
<tr>
<td>Other</td>
<td>3,102 (13.5)</td>
</tr>
<tr>
<td>Gender</td>
<td></td>
</tr>
<tr>
<td>Boys</td>
<td>10,954 (47.8)</td>
</tr>
<tr>
<td>Girls</td>
<td>11,971 (52.2)</td>
</tr>
</tbody>
</table>

a The OHP administers both Oregon’s Medicaid Program and CHIP.
b Child had an outpatient visit in 2010 at 1 of the 46 Oregon clinics that had fully implemented OCHIN’s EHR before July 1, 2009.
information can be found in claims and EHR data; thus, accurate assessment of care quality may require combining these data sources. The “hybrid” method recommended for assessing CHIPRA quality measures, which includes medical record review, is likely more comprehensive than analyses that use administrative claims data alone. However, this method, as currently described, is difficult to execute because of time and financial constraints.

We demonstrate an alternative data-collection method for those systems able to electronically abstract clinically relevant EHR data; this method may improve the ability to use medical record data for quality assessment.

Impact on Policy

The Medicaid Statistical Information System data, a national eligibility and claims database maintained by the Centers for Medicare and Medicaid Services, has been identified as a potential way to further standardize the CHIPRA core measures as states expand coverage by enacting provisions authorized under the Affordable Care Act.25,26 Using Medicaid Statistical Information System data will allow for state-to-state comparison, as measures will be calculated the same way from a national dataset. However, our results suggest that these data alone may yield an inaccurate measure of pediatric care quality.

Policy makers should consider combining claims and EHR data to more accurately assess the quality of care provided to children insured by Medicaid and CHIP. As reimbursements are increasingly based on the quality of delivered care, it will be essential to access and use the most complete data available.27 As more and more health care providers have EHRs,28 health record data will be increasingly available.
to supplement claims data. With expanded availability comes the need for improved methods for accessing and using EHR data. The data must be organized, synthesized, and easily aggregated and obtained from the EHR's "back end," and standardized data entry processes are needed on the "front end" to ensure that EHR information is as systematic and complete as possible. The end user also needs to be considered when making future EHR improvements. For example, there are differences in pediatric patient care that should be reflected in the front end to ensure the appropriate data are recorded. Additionally, the increasing use and development of EHR data also suggest the need to change the pediatric quality measurement technical specifications to include automated extraction of data from discrete EHR data fields, in addition to (or in place of) a manual chart review hybrid method. Extracting EHR data automatically has been shown to perform reasonably well, compared with manual chart review. The transition from International Classification of Diseases, Ninth Revision (ICD-9) diagnosis codes to International Classification of Diseases, Tenth Revision, Clinical Modification (ICD-10-CM) codes may help improve the accuracy and completeness of claims data, as ICD-10-CM is a more robust coding system with up to 7 characters to accommodate new treatments and technologies and performance-based payment systems. The use of ICD-10-CM codes coupled with billing practices that align closely with the technical specifications of quality metrics may improve the accuracy of claims data for quality reporting. However, as payment moves away from fee-for-service toward capitation and outcome-based global reimbursement, claims-based quality measurement may become obsolete, and quality measurement that uses clinical data from EHRs may become even more essential.

This study demonstrates how EHRs are poised to transform the way health care is delivered, as well as how health services and health care delivery can be studied. Significant future investments will be required to further develop and most effectively use EHRs to fully maximize potential gains.

The examples presented in this article illustrate how measuring health care quality can vary depending on the data source(s) used. The collection and storage of EHR data differs between care provider organizations, thus, including additional EHR fields or assessing EHR data from other systems could yield different results. Additionally, OCHIN's centrally hosted EHR data are regularly checked and cleaned; these practices are not standard across all health care organizations. We limited these analyses to children continuously insured by the Oregon Health Plan who were established patients in 1 of the 46 Oregon clinics that had implemented a full EHR before July 1, 2009. Similar analyses may yield different results if conducted in other states or clinics; our purpose in studying this specific cohort of children was to assess the same measures in the same population calculated by using different data sources.

CONCLUSIONS

Children's care quality measures assessed by using administrative claims alone may not be accurate. Adding EHR data to administrative claims data may yield more complete measurement.

REFERENCES


27. Frakt AB, Mayes R. Beyond capitation: how new payment experiments seek to find the ‘sweet spot’ in amount of risk providers and payers bear. *Health Aff (Millwood)*. 2012;31(9):1951–1958


Variation in Outcomes of Quality Measurement by Data Source
Heather Angier, Rachel Gold, Charles Gallia, Allison Casciato, Carrie J. Tillotson, Miguel Marino, Rita Mangione-Smith and Jennifer E. DeVoe

Pediatrics 2014;133:e1676; originally published online May 26, 2014; DOI: 10.1542/peds.2013-4277

Updated Information & Services
including high resolution figures, can be found at:
/content/133/6/e1676.full.html

References
This article cites 18 articles, 8 of which can be accessed free at:
/content/133/6/e1676.full.html#ref-list-1

Citations
This article has been cited by 1 HighWire-hosted articles:
/content/133/6/e1676.full.html#related-urls

Subspecialty Collections
This article, along with others on similar topics, appears in the following collection(s):
Health Information Technology
/cgi/collection/health_information_technology_sub
Electronic Health Records
/cgi/collection/electronic_health_records_sub
Advocacy
/cgi/collection/advocacy_sub
Federal Policy
/cgi/collection/federal_policy_sub

Permissions & Licensing
Information about reproducing this article in parts (figures, tables) or in its entirety can be found online at:
/site/misc/Permissions.xhtml

Reprints
Information about ordering reprints can be found online:
/site/misc/reprints.xhtml

PEDIATRICS is the official journal of the American Academy of Pediatrics. A monthly publication, it has been published continuously since 1948. PEDIATRICS is owned, published, and trademarked by the American Academy of Pediatrics, 141 Northwest Point Boulevard, Elk Grove Village, Illinois, 60007. Copyright © 2014 by the American Academy of Pediatrics. All rights reserved. Print ISSN: 0031-4005. Online ISSN: 1098-4275.
Variation in Outcomes of Quality Measurement by Data Source
Heather Angier, Rachel Gold, Charles Gallia, Allison Casciato, Carrie J. Tillotson, Miguel Marino, Rita Mangione-Smith and Jennifer E. DeVoe
Pediatrics 2014;133:e1676; originally published online May 26, 2014;
DOI: 10.1542/peds.2013-4277

The online version of this article, along with updated information and services, is located on the World Wide Web at:
/content/133/6/e1676.full.html