Patient-Centered Outcomes Measurement: Does It Require Information From Patients?

Leif I. Solberg
Stephen E. Asche
John Butler
David Carrell
Christine K. Norton
Jeffrey G. Jarvik
Rebecca Smith-Bindman
Juliana O. Tillema
Robin R. Whitebird
Ann M. Werner

See next page for additional authors

Follow this and additional works at: https://aurora.org/jpcrr

Part of the Health Information Technology Commons, and the Health Services Research Commons

Recommended Citation
Patient-Centered Outcomes Measurement: Does It Require Information From Patients?

Authors
Patient-Centered Outcomes Measurement: Does It Require Information From Patients?

Leif I. Solberg, MD,1 Stephen E. Asche, MA,1 John Butler, MD,2 David Carrell, PhD,3 Christine K. Norton, MA,4 Jeffrey G. Jarvik, MD, MPH,4 Rebecca Smith-Bindman, MD,5 Juliana O. Tillema, MPA,1 Robin R. Whitebird, PhD,6 Ann M. Werner, BS,1 Jeanette Y. Ziegenfuss, PhD1

1HealthPartners Institute for Education and Research, Minneapolis, MN; 2HealthPartners Medical Group, Minneapolis, MN; 3Group Health Research Institute, Seattle, WA; 4University of Washington, Seattle, WA; 5University of California, San Francisco, San Francisco, CA; 6University of St. Thomas, St. Paul, MN; 7No affiliation (patient researcher)

Purpose
Since collecting outcome measure data from patients can be expensive, time-consuming, and subject to memory and nonresponse bias, we sought to learn whether outcomes important to patients can be obtained from data in the electronic health record (EHR) or health insurance claims.

Methods
We previously identified 21 outcomes rated important by patients who had advanced imaging tests for back or abdominal pain. Telephone surveys about experiencing those outcomes 1 year after their test from 321 people consenting to use of their medical record and claims data were compared with audits of the participants’ EHR progress notes over the time period between the imaging test and survey completion. We also compared survey data with algorithmically extracted data from claims files for outcomes for which data might be available from that source.

Results
Of the 16 outcomes for which patients’ survey responses were considered to be the best information source, only 2 outcomes for back pain and 3 for abdominal pain had kappa scores above a very modest level of ≥ 0.2 for chart audit of EHR data and none for algorithmically obtained EHR/claims data. Of the other 5 outcomes for which claims data were considered to be the best information source, only 2 outcomes from patient surveys and 3 outcomes from chart audits had kappa scores ≥ 0.2.

Conclusions
For the types of outcomes studied here, medical record or claims data do not provide an adequate source of information except for a few outcomes where patient reports may be less accurate. (J Patient Cent Res Rev. 2017;4:221-229.)

Keywords
electronic health records, patient-reported outcomes, patient-centered care
The objective of this study was to answer that question as well as learn the degree of agreement among different sources of outcome information.

**METHODS**

The methods used to identify and prioritize outcomes important to patients who had back or abdominal pain significant enough to have their primary care physicians order advanced imaging studies (computerized tomography or magnetic resonance imaging) have been previously described. First, outcomes of importance were identified by interviews with patients and family members, after which 40 patients and 8 family members rated the importance of each outcome. The 21 outcomes are listed in Table 1 in rank order of importance to patients on a scale of 1–5, wherein 5 equals extremely important and 1 equals not at all important.

We then conducted a telephone survey of 368 adult patients, 12–20 months following their imaging study for back or abdominal pain, who had both insurance and

<table>
<thead>
<tr>
<th>Outcome</th>
<th>Mean</th>
<th>SD</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. To find out the cause of the pain</td>
<td>4.9</td>
<td>0.4</td>
</tr>
<tr>
<td>2. To trust that the treatment plan is appropriate</td>
<td>4.7</td>
<td>0.6</td>
</tr>
<tr>
<td>3. To return to normal life functions</td>
<td>4.7</td>
<td>0.6</td>
</tr>
<tr>
<td>4. Satisfied with results of care</td>
<td>4.7</td>
<td>0.8</td>
</tr>
<tr>
<td>5. To understand what may happen to you because of the problem</td>
<td>4.6</td>
<td>0.7</td>
</tr>
<tr>
<td>6. To prevent this problem from occurring again</td>
<td>4.6</td>
<td>0.9</td>
</tr>
<tr>
<td>7. To prevent long-term loss of function</td>
<td>4.6</td>
<td>0.9</td>
</tr>
<tr>
<td>8. To return to work and productivity as soon as possible</td>
<td>4.5</td>
<td>0.9</td>
</tr>
<tr>
<td>9. Satisfied with how care delivered</td>
<td>4.5</td>
<td>0.9</td>
</tr>
<tr>
<td>10. To experience no complications or side effects</td>
<td>4.3</td>
<td>1.0</td>
</tr>
<tr>
<td>11. To be assured that no unexpected, unrelated problems develop</td>
<td>4.2</td>
<td>1.2</td>
</tr>
<tr>
<td>12. To get rapid and complete relief from pain and other symptoms</td>
<td>4.2</td>
<td>1.0</td>
</tr>
<tr>
<td>13. To avoid being hospitalized</td>
<td>4.1</td>
<td>1.2</td>
</tr>
<tr>
<td>14. To avoid surgery</td>
<td>4.1</td>
<td>1.3</td>
</tr>
<tr>
<td>15. To avoid placing a burden or stress on family members</td>
<td>4.1</td>
<td>1.1</td>
</tr>
<tr>
<td>16. To minimize or avoid the need for further tests and medical visits</td>
<td>3.9</td>
<td>1.1</td>
</tr>
<tr>
<td>17. To minimize radiation exposure in the course of my care</td>
<td>3.9</td>
<td>1.4</td>
</tr>
<tr>
<td>18. To avoid personal costs for care</td>
<td>3.9</td>
<td>1.4</td>
</tr>
<tr>
<td>19. To minimize or avoid use of medication</td>
<td>3.8</td>
<td>1.1</td>
</tr>
<tr>
<td>20. To return to leisure/sports activities as soon as possible</td>
<td>3.8</td>
<td>1.1</td>
</tr>
<tr>
<td>21. To minimize discomfort from the tests used to assess the pain</td>
<td>3.7</td>
<td>1.3</td>
</tr>
</tbody>
</table>

*1 = not at all important, 5 = extremely important. SD, standard deviation.
care from a single large health care system so complete information would be available. Since existing validated instruments for collecting patient-reported outcomes do not address most of the outcomes patients told us they preferred, we needed to create and pretest our own unvalidated survey. The survey asked about whether patients had experienced each of the outcomes and, if so, when they did, since many of the outcomes related to time period of occurrence. The response rate for survey completion among those potentially eligible was 56% (368 respondents out of 656 invited); 67% of the 552 who we were able to contact completed the survey.

Of these 368 patients, 321 (87.2%, 151 with back pain and 170 with abdominal pain) also agreed to allow the research team to review their medical records and health insurance claims information during the time interval between the imaging study and survey completion. We tested two different methods for searching for information on these outcomes within those databases:

1. Manual chart audit of unstructured progress notes in the EHR (henceforth called “chart audit”) by experienced medical auditors trained to identify information relevant to 20 specific outcomes. The care system studied has transitioned all medical record information to an electronic form, so there was no need to search paper records for this review. The outcome of returning to normal life functions was dropped a priori as not determinable from progress notes, yielding a total of 20 outcomes for manual chart audit. Two trained medical auditors reviewed EHR progress note data for all visits between the time of the index scan and survey completion in which the visit type was phone, internet or in-person with a physician, nurse practitioner, physician assistant, nurse or physical therapist (for back pain) in a potentially relevant department. Auditors were provided with definitions of what constituted each outcome (eg, what timing for pain relief could be considered rapid), and searched all visit data for evidence of the outcome. If the outcome was coded as “definitely found” in any patient visit, that patient was coded as having the outcome occurring. Rater agreement for the last 16 patients had a mean kappa of 0.66 across the 20 items assessed by chart audit.

2. Use of algorithmic programs to search for discrete data in structured fields within medical record or claims data (henceforth called “discrete data”) about any of the outcomes with the potential for containing information relevant to the outcome (eg, hospitalization, tests, surgery). The algorithms were developed by several of the authors in an iterative approach with repeated testing and revision. However, for 9 of these outcomes, no terms or combination of terms was successful in finding relevant information in either the EHR or claims, since these kinds of patient desires are not part of either record. These searches focused on standard structured fields and did not include structured elements like smart notes that some clinicians might use within their progress notes because the study goal was to learn whether easily accessible structured information might be helpful in identifying outcome information.

Assessment of Data Sources
To assess the relative value of the three data sources (patient survey, chart or discrete data) the entire author group reviewed our data on the frequency of outcome information and made the following decisions about how to conduct the analysis:

- For those 5 outcomes with specific medical services that were covered by health insurance and would normally have an associated claim for payment, the discrete data algorithmically pulled from the EHR and claims were assumed to be the most accurate source of information about whether a service was provided and when. This included hospitalizations, surgery, tests and medical visits, imaging exams, and prescription medications. In each case, this information would also be present in a discrete format in the medical record unless it was received through a medical provider who was outside the study health system. However, the algorithm credited the service if it was present in either source to ensure comprehensiveness. If these data disagreed with patient report about the event, we considered the record/claims source to be correct since patients might not recall the occurrence or timing of these specific objective events accurately.

- For the other 16 outcomes, the patient was assumed to be the most accurate source of information. Although there might be problems with recall, each of these outcomes has a large subjective element (eg, pain relief, confidence, satisfaction) or involved information that could only be known and defined by the patient (eg, return to life functions and activities,
family burden, personal costs). Therefore, information from the patient about these outcomes was chosen as the preferred source of these data. Indeed, for only 5 of these outcomes did we believe it was possible to even consider examining algorithm-based discrete data sources.

Given these assumptions, the important analytic questions became: A) For the 16 outcomes for which the patient was the standard, could discrete data from the medical record or claims provide an acceptable substitute?; and B) For the 5 objective outcomes for which algorithmically derived discrete data were the standard, how closely did patient surveys or chart audit of the unstructured progress notes mirror those findings?

All aspects of this study were reviewed, approved and monitored by the local institutional review board.

Analysis
Mean importance ratings of outcomes were summarized using data from a prior study. Demographic characteristics of the 321 survey respondents are presented separately for abdominal pain and back pain patients and presented in the form of counts and percentages. Differences in patient attributes by pain type are tested with independent sample t-tests and Fisher’s exact test. The percentage of abdominal pain and back pain patients having each outcome is summarized separately by data sources of patient survey, discrete data (algorithmic pull of EHR/claims data), and manual chart audit of the unstructured progress notes in the EHR.

For the 16 outcomes thought to be ideally collected by patient survey, the agreement on the presence of the outcome between the patient survey and both the discrete data and chart audit data was summarized with Cohen’s kappa (and 95% confidence interval), sensitivity, specificity, and positive and negative predictive value. For the 5 outcomes thought to be ideally obtained from discrete data, the agreement on the presence of the outcome between the discrete data and both the patient survey and chart audit data was summarized using the same summary statistics. We limited the presentation of outcomes to those with kappa ≥ 0.2 to focus summaries on outcomes with at least a minimal level of concordance across data sources.14

RESULTS
The characteristics of the 321 survey respondents who also agreed to allow review of their medical records and claims data are listed in Table 2. Overall, these patients were mostly white (88%), middle-aged, female, and well-educated, although only 57% were employed. Surprisingly, there were no statistically significant differences in demographic variables between people with the two pain types.

Table 3 compares the frequency of each outcome (separately for back and abdominal pain patients) as reported by patient surveys, discrete data or chart audits. Note that those outcomes patients wish to avoid (recurrence, function loss, burden, costs, surgery, etc.) are reworded instead as positive statements since listing them in negative terms incorrectly suggests a greater ability to find information about these outcomes in alternate sources. Data are displayed in two parts to facilitate comparisons between what we suggest is the best comparison and the other two feasible alternatives. It is readily apparent for those outcomes dependent on the patient survey that only 5 of the 16 outcomes have the possibility of a discrete data comparison, and there is little correspondence between chart audit or discrete data and patient survey answers. Out of the 10 most important outcomes, patients appear to have largely achieved most of them. The main disappointments appear to have been “Learned the cause” (abdominal pain), “Returned to normal function” and “Return to work” soon (both types), and “Had treatment side effects” (back pain). For most outcomes, the chart audit contained far less information for identifying outcomes received than the patient survey. It is only comparable for “Learned the cause” and “Learned what may happen,” whereas it greatly overestimates the likelihood of “Obtained rapid and complete symptom relief.”

In outcomes for which discrete data seemed to be the best source, all three sources are much more similar in their estimates of obtaining desired outcomes. Patients especially underestimated having had further tests/visits and additional radiation exposures (except for abdominal pain patients) than the chart audits or discrete data demonstrated. They also were more likely to identify avoiding medications than the other sources.
Finally, in Tables 4 and 5, the degree of agreement (kappa) between specific data sources for each outcome are shown for those outcomes in which the agreement was at least fair (kappa ≥ 0.2). For those outcomes, we also show the sensitivity, specificity and predictive values. Only 3 chart audit outcomes (“Learned the cause,” “Had treatment side effects,” and “Had unexpected problems”) met this agreement criterion for an alternate source of information for items for which the patient survey response was considered the standard (Table 4). However, 2 of the 5 outcomes (“Hospitalization” and “Surgery”) with a discrete data standard (Table 5) met this criterion for both patient survey and chart audit.

**DISCUSSION**

Although these results are limited to two specific conditions (back and abdominal pain) and a particular setting, they suggest that alternate sources of data about these patient outcomes are quite limited. Only 2 of the patient-preferred outcomes can be approximated by chart audits for both pain types (“Learned the cause” and “Had unexpected problems”), and only 2 discrete data outcomes can be approximated by either patient surveys or chart audits (“Had hospitalization” and “Had surgery”). In most cases, the specificity and negative predictive value are much better than the sensitivity and positive predictive value. In other words, the medical sources appear to miss a fair number of outcomes. Thus, both clinicians and researchers need to obtain information about these types of outcomes from patients and would not usually be able to find this information in more easily accessible medical sources. This also means that outcomes important to patients are not being documented and likely not being discussed.
so clinicians and care systems often wouldn’t know about poor outcomes from the patients’ perspectives unless they asked directly.

Some of the discordances between the patient survey and our medical data sources probably reflect varying perceptions, values and recall by individual patients. However, there is also a technical reason contributing to these differences. For both manual chart audits and algorithmically derived discrete data, it is necessary to set specific cutoff points for time periods, numbers of events and types of services. These chosen time points may often be different than the cutoffs used by individual patients as they manage and deal with their symptoms. Finally, while each of these outcomes had high average importance ratings, there will always be some people for whom particular outcomes are much more or much less important, causing them to respond to surveys about those outcomes in ways that create differences from our standard quantitative approach.

Since the field of patient-reported outcome measurement is still in its early stages, these findings are difficult to compare with the limited literature in this area. One of the few source studies is that by Gresham et al, who found fairly good agreement between self-reported perinatal outcomes and medical records.15 However, there is substantial literature on alternate sources of information about a variety of traditional clinical information and events. For example, studies

---

**Table 3. Frequency of Patient-Desired Outcomes by Data Source Documenting It**

<table>
<thead>
<tr>
<th>Outcome*</th>
<th>Back pain (N=151)</th>
<th>Abdominal pain (N=170)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Survey</td>
<td>Algorithm</td>
</tr>
<tr>
<td>Learned cause</td>
<td>80%</td>
<td>X 54%</td>
</tr>
<tr>
<td>Treatment plan appropriate</td>
<td>76%</td>
<td>X 24%</td>
</tr>
<tr>
<td>Return to normal function</td>
<td>30%</td>
<td>X 15%</td>
</tr>
<tr>
<td>Satisfied with results</td>
<td>74%</td>
<td>X 0%</td>
</tr>
<tr>
<td>Learn what may happen</td>
<td>79%</td>
<td>X 93%</td>
</tr>
<tr>
<td>Had recurrence</td>
<td>8%</td>
<td>68% 25%</td>
</tr>
<tr>
<td>Had function loss</td>
<td>64%</td>
<td>X 2%</td>
</tr>
<tr>
<td>Return to work, productivity</td>
<td>7%</td>
<td>X 5%</td>
</tr>
<tr>
<td>Satisfied with care</td>
<td>79%</td>
<td>X 0%</td>
</tr>
<tr>
<td>Had side effects</td>
<td>46%</td>
<td>7% 3%</td>
</tr>
<tr>
<td>Had unexpected problems</td>
<td>5%</td>
<td>21% 5%</td>
</tr>
<tr>
<td>Got rapid/complete relief</td>
<td>5%</td>
<td>16% 42%</td>
</tr>
<tr>
<td>Burdened others</td>
<td>52%</td>
<td>X 0%</td>
</tr>
<tr>
<td>Had personal costs</td>
<td>48%</td>
<td>X 1%</td>
</tr>
<tr>
<td>Returned to leisure/sports</td>
<td>29%</td>
<td>X 1%</td>
</tr>
<tr>
<td>Had test discomfort</td>
<td>40%</td>
<td>4% 0%</td>
</tr>
<tr>
<td>Had hospitalization</td>
<td>7%</td>
<td>15% 8%</td>
</tr>
<tr>
<td>Had surgery</td>
<td>13%</td>
<td>14% 11%</td>
</tr>
<tr>
<td>Had further tests, visits</td>
<td>51%</td>
<td>91% 95%</td>
</tr>
<tr>
<td>Had additional radiation</td>
<td>3%</td>
<td>54% 23%</td>
</tr>
<tr>
<td>Had considerable meds</td>
<td>61%</td>
<td>78% 83%</td>
</tr>
</tbody>
</table>

*Items are sorted by importance with the most important at the top. Undesired outcomes are worded differently than in Table 1 to avoid the false impression that data on the outcome were available most of the time.

X = data not available.
have found good concordance between patient surveys and either EHR or claims data for identifying whether patients had certain diagnoses (diabetes, bowel disease), but not for others (hypertension, cardiovascular disease).\textsuperscript{16-18} Gupta et al also found good concordance between patient self-report and the medical record on a variety of events and treatments for breast cancer.\textsuperscript{19} Tisnado et al measured concordance between the medical record and patient report in ambulatory care and concluded that it varied by domain — good for diagnosis, clinical services and medication use, and poor for counseling and referrals.\textsuperscript{20} However, there was considerable variation in agreement within these domains, depending on the specific items.

Many studies have focused on the accuracy of patient reports of utilization as compared to administrative claims data and generally have found good agreement

\begin{table}[h]
\centering
\caption{Chart Audit Agreement Metrics\textsuperset{a} for Significant Outcomes With a Patient Survey Standard}
\begin{tabular}{|l|c|c|c|c|c|c|c|}
\hline
Outcome & Survey & Audit & Kappa (95\% CI) & Sens & PPV & Spec & NPV \\
\hline
Back pain & & & & & & & \\
Learned cause & 80\% & 54\% & 0.20 (0.06–0.33) & 0.60 & 0.89 & 0.70 & 0.30 \\
Had unexpected problems & 5\% & 5\% & 0.21 (-0.08–0.50) & 0.25 & 0.25 & 0.96 & 0.96 \\
\hline
Abdominal pain & & & & & & & \\
Learned cause & 65\% & 29\% & 0.21 (0.10–0.31) & 0.38 & 0.86 & 0.88 & 0.43 \\
Had side effects & 9\% & 2\% & 0.20 (-0.05–0.45) & 0.13 & 0.67 & 0.99 & 0.92 \\
Had unexpected problems & 12\% & 18\% & 0.20 (0.02–0.38) & 0.38 & 0.27 & 0.85 & 0.91 \\
\hline
\end{tabular}
\textsuperscript{a}Items selected and sorted by kappa. Only kappa \( \geq 0.2 \) retained.
\end{table}

\begin{table}[h]
\centering
\caption{Patient Survey and Chart Audit Agreement Metrics\textsuperset{a} for Significant Outcomes With an Algorithmic Data Standard}
\begin{tabular}{|l|c|c|c|c|c|c|c|}
\hline
Outcome & Algorithm & Survey & Kappa (95\% CI) & Sens & PPV & Spec & NPV \\
\hline
Back pain & & & & & & & \\
Had hospitalization & 15\% & 7\% & 0.23 (0.02–0.44) & 0.23 & 0.45 & 0.95 & 0.88 \\
Had surgery & 14\% & 13\% & 0.42 (0.21–0.63) & 0.48 & 0.53 & 0.93 & 0.92 \\
\hline
Abdominal pain & & & & & & & \\
Had hospitalization & 8\% & 11\% & 0.24 (0.02–0.46) & 0.38 & 0.26 & 0.91 & 0.95 \\
Had surgery & 18\% & 17\% & 0.49 (0.31–0.66) & 0.57 & 0.59 & 0.91 & 0.91 \\
\hline
\end{tabular}
\textsuperscript{a}Items selected and sorted by kappa. Only kappa \( \geq 0.2 \) retained.
\end{table}

\begin{table}[h]
\centering
\caption{Comparison of Algorithmic Data and Survey}
\begin{tabular}{|l|c|c|c|c|c|c|c|}
\hline
Outcome & Algorithm & Survey & Kappa (95\% CI) & Sens & PPV & Spec & NPV \\
\hline
Back pain & & & & & & & \\
Had hospitalization & 15\% & 8\% & 0.34 (0.13–0.56) & 0.32 & 0.58 & 0.96 & 0.89 \\
Had surgery & 14\% & 11\% & 0.42 (0.20–0.63) & 0.43 & 0.56 & 0.95 & 0.91 \\
Had radiation tests & 54\% & 23\% & 0.30 (0.18–0.42) & 0.37 & 0.88 & 0.94 & 0.56 \\
\hline
Abdominal pain & & & & & & & \\
Had hospitalization & 8\% & 14\% & 0.38 (0.17–0.60) & 0.62 & 0.35 & 0.90 & 0.97 \\
Had surgery & 18\% & 14\% & 0.53 (0.36–0.71) & 0.53 & 0.70 & 0.95 & 0.90 \\
\hline
\end{tabular}
\textsuperscript{a}Items selected and sorted by kappa. Only kappa \( \geq 0.2 \) retained.
\end{table}

\begin{table}[h]
\centering
\caption{Comparison of Algorithmic Data and Chart Audit}
\begin{tabular}{|l|c|c|c|c|c|c|c|}
\hline
Outcome & Algorithm & Audit & Kappa (95\% CI) & Sens & PPV & Spec & NPV \\
\hline
Back pain & & & & & & & \\
Had hospitalization & 15\% & 8\% & 0.34 (0.13–0.56) & 0.32 & 0.58 & 0.96 & 0.89 \\
Had surgery & 14\% & 11\% & 0.42 (0.20–0.63) & 0.43 & 0.56 & 0.95 & 0.91 \\
Had radiation tests & 54\% & 23\% & 0.30 (0.18–0.42) & 0.37 & 0.88 & 0.94 & 0.56 \\
\hline
Abdominal pain & & & & & & & \\
Had hospitalization & 8\% & 14\% & 0.38 (0.17–0.60) & 0.62 & 0.35 & 0.90 & 0.97 \\
Had surgery & 18\% & 14\% & 0.53 (0.36–0.71) & 0.53 & 0.70 & 0.95 & 0.90 \\
\hline
\end{tabular}
\textsuperscript{a}Items selected and sorted by kappa. Only kappa \( \geq 0.2 \) retained.
\end{table}
for whether a service was performed.21-26 However, for many of the items we studied, the timing of an outcome was as important as whether it occurred, and that has not been reported in any studies to our knowledge, other than some studies showing that patients tend to think screening tests occurred more recently than was the case.27-29 Taken collectively, these and other studies suggest that while agreement between various sources of information is fairly good on average, there is considerable variation by type of information and need for details.

As suggested earlier, this study has important limitations, both in its technical approach (newly developed unvalidated survey) and its limit to a particular setting with a low proportion of nonwhite patients and two specific clinical situations with heterogeneous causes. It is also the first time we are aware of that these types of mostly subjective outcomes have been measured and compared. Thus, our results are preliminary findings that need to be tested by other studies. Most of the attention currently directed toward measuring and reporting patient-reported outcomes has involved the intentionally quantitative symptom and function outcomes in NIH’s PROMIS system.5,30 Valuable as that approach might be, our experience with this project has suggested that other subjective outcomes are also important to patients, and if we want to become more patient-centered in medical care, we shall need to learn how to collect, report and use information about these other patient outcomes as well.

As an important sidelight from this study, we have clearly learned that many of the outcomes important to patients are not currently being documented in medical records. This may be a missed opportunity in our efforts to be more patient-centered. A systematic approach to collecting, recording and using these important data is needed. As we develop such systems for new types of information important to patients, it should be helpful to have studies like this as confirmation that there are few good substitutes for asking patients, both about what matters to them and whether it occurred.

**CONCLUSIONS**

This study demonstrates that most outcome information important to patients cannot be obtained from electronic medical records or insurance claims data, therefore requiring such information be obtained directly from patients, either during clinical encounters or via mail, telephone or electronic surveys. With such data, patients can be informed about how their outcomes compare with similar patients in order to make adjustments in management and to compare care among clinicians for purposes of guiding selection of clinicians and quality improvement.

**Patient-Friendly Recap**

- Outcome data that are most important to patients could assist them in selecting clinicians, predicting likelihood of achieving desired outcomes, and indicating the need for changes in care.
- However, this information is currently infrequently available in existing medical records or insurance claims systems.
- Therefore, patient-friendly methods of obtaining this information directly from patients must be developed and used widely throughout the care process.

**Conflicts of Interest**

None.

**Funding Sources**

This work was supported through a Patient-Centered Outcomes Research Institute (PCORI) Pilot Project Program Award. Disclaimer: All statements in this report, including its findings and conclusions, are solely those of the authors and do not necessarily represent the views of PCORI, its Board of Governors, or Methodology Committee.

**References**


12. Landis JR, Koch GG. The measurement of observer agreement for categorical data. *Biometrics.* 1977;33:159-74. CrossRef


© 2017 Aurora Health Care, Inc.