Keen Eye on Core Measures: Joint Commission Data Quality Study Offers Insights into Data Collection, Abstracting Processes

by Ann Watt, MBA, RHIA; Scott Williams, PsyD; Kwan lee, PhD; James Robertson; Richard D. Koss, MA; and Jerod M. Loeb, PhD

The Joint Commission has been collecting core measures data for more than a year. But how reliable are the data being collected? Strategies from successful organizations show how improving your documentation processes can improve the reporting of core measures.

In July 2002, the Joint Commission on Accreditation of Healthcare Organizations implemented a requirement that accredited US hospitals collect and report data on evidence-based, standardized performance measures. Four core measure sets were developed: acute myocardial infarction (AMI), heart failure (HF), community-acquired pneumonia (CAP), and pregnancy and related conditions (PR).

Since then, the majority of Joint Commission-accredited hospitals have been required to report on two of these initial core measure sets based on the services they provide. (This requirement has been increased to three core measure sets beginning in January 2004.) Because this process involves collecting and abstracting data, it frequently has an impact on HIM professionals.

Establishing and maintaining the quality of national core performance measure data are critical because the data will be used for a variety of purposes, including quality improvement and public reporting. To demonstrate the impact of evidence-based performance measurement on healthcare quality, the Joint Commission is conducting a research project designed to assess the completeness and accuracy of the data flowing into the national comparative core measures database and to evaluate improvement actions taken by healthcare organizations.

This article summarizes results to date and offers some suggestions for resolving documentation issues to improve core measures data.

How the Study Worked

The initial phase of this project, conducted during the first half of 2003, involved on-site visits by Joint Commission staff members to 30 randomly selected participating hospitals to assess the inter-rater reliability of different abstractors collecting the same core measure.

Hospital participation in this activity was voluntary, and participating organizations were stratified by geographical area, number of beds, urban or rural location, and for-profit or not-for-profit status.

During the site visits, Joint Commission staff abstracted core measure data elements previously submitted by the hospitals to the national core measures database from approximately 30 medical records (15 records per measure set). These reabstracted data were then compared with the original data and any discrepancies noted were discussed with hospital staff to determine the reasons for the differences.

To determine the degree to which discrepancies in individual data elements impacted measure rates, the original and reabstracted records were each processed through the core measure algorithms. The resulting rates—one derived from the

original abstraction by the hospital and the second resulting from reabstraction by Joint Commission staff—were then compared to assess level of agreement.

These agreement rates for each measure can be used as an indicator of overall database reliability. "Measure Agreement Rates," below, includes the agreement rates and Kappa statistics for each core measure.

Measure Agreement Rates

Measure	Measure Name	n	Agreement Rate	Kappa
AMI 1	Aspirin at arrival	227	0.84	0.68
AMI 2	Aspirin prescribed at discharge	227	0.82	0.69
AMI 3	ACEI for patients with LVSD	227	0.94	0.66
AMI 4	Adult smoking cessation advice/counseling	227	0.97	0.87
AMI 5	Beta blocker prescribed at discharge	227	0.90	0.84
AMI 6	Beta blocker at arrival	227	0.80	0.67
AMI 7	Time to thrombolysis	227	0.91	0.51
AMI 8	Time to PTCA	227	0.90	0.44
AMI 9	Inpatient mortality	227	0.96	0.92
HF 1	Discharge instructions	217	0.88	0.81
HF 2	Left ventricular function assessment	217	0.88	0.74
HF 3	ACEI for patients with LVSD	217	0.89	0.73
HF 4	Adult smoking cessation advice/counseling	217	0.94	0.78
CAP 1	Oxygenation assessment	202	0.83	0.50
CAP 2	Pneumococcal screening and/or vaccination	202	0.92	0.86
CAP 3	Blood cultures	202	0.77	0.60
CAP 4a	Adult smoking cessation advice/counseling	202	0.91	0.65
CAP 4b	Pediatric smoking cessation advice/counseling	202	0.99	0.66
CAP 5	Antibiotic timing	202	0.78	0.43
PR 1	VBAC	118	0.99	0.91
PR 2	Inpatient neonatal mortality	118	1.00	1.00
PR 3	Third or fourth degree laceration	118	0.99	0.98

The Kappa statistic represents the agreement rate between outcomes, after the probability of chance agreement has been removed. Kappa values that are greater than 0.75 are generally considered to be representative of excellent agreement beyond chance. Values between 0.40 and 0.75 represent fair to good agreement beyond chance, and values below 0.40 represent poor agreement.[1] The total number of reabstracted records is represented by the "n" column.

Reliability of individual data elements was also assessed. "Data Element Agreement Rates," below, presents a summary of data element reliability by measure set. Regardless of which measure sets were selected by a hospital, certain general data

elements must be collected and submitted for every patient that falls into any of the selected measure set populations. Measure-specific data elements are those data elements that are linked to a particular measure and are usually collected through medical record reabstraction.

Data Element Agreement Rates

	Mean Agreement Rate	Mean Kappa	Number of Data Elements
AMI General Data Elements	0.955	0.899	11
AMI Specific Data Elements	0.914	0.644	22
AMI All Data Elements	0.928	0.713	33
HF General Data Elements	0.932	0.889	11
HF Specific Data Elements	0.907	0.700	13
HF All Data Elements	0.919	0.770	24
CAP General Data Elements	0.959	0.937	11
CAP Specific Data Elements	0.912	0.745	18
CAP All Data Elements	0.930	0.818	29
PR General Data Elements	0.945	0.894	11
PR All Data Elements	0.941	0.879	12
General data elements include patient demographic information such as case identifier, sex and birth date; visit-specific data including admission and discharge dates; ICD-9-CM codes; and detail as to admission type and source and discharge disposition.			

What the Study Found: General and Specific Issues

The data quality issues identified during the reabstraction site visits can be categorized into those involving general data elements, which are common to all measure sets, and those involving measure-specific clinical data elements.

General Data Elements

In many cases, because of their administrative nature, general data element values are electronically transferred from a healthcare organization's UB-92 billing database into its core measures abstraction database. Site visit reabstraction findings indicated that discrepancies sometimes occurred between the transferred administrative data and the data contained in the medical record.

For example, in a number of cases, originally abstracted data (i.e., pre-populated from the hospital's billing database) indicated that patients were discharged to home. Upon reabstraction, review of the medical record revealed that the patients actually were transferred to another hospital or to a skilled nursing facility.

This discrepancy can have a significant impact on the determination as to whether a patient is even included in the population for certain performance measures and thus can significantly impact the healthcare organization's reported performance rate for that measure. Similarly, discrepancies in admission source can have the same result. Therefore, to promote core measures database reliability, abstractors should confirm the accuracy of pre-populated data and have the ability to override it if it is inaccurate.

Measure-Specific Data Elements

Acute Myocardial Infarction: The data element from this measure set exhibiting the greatest degree of discrepancy between the original abstraction and site visitors' reabstraction was the one related to initial electrocardiogram (ECG)

interpretation. Dialogue with hospital abstraction staff indicated that the reasons for this disagreement ranged from confusion on the part of abstractors regarding interpretation of the data element definition to amendment of the definition by hospital clinical staff to unclear documentation in the medical record.

In some cases, hospital staff, apparently under the misimpression that the data definitions and abstraction instructions were only suggested guidelines, open to interpretation, expanded the definition for the abstraction instructions to include ECG interpretations that should have been excluded. Problems with correct abstraction of this data element also arose on occasion when cardiologists interpreting the ECGs provided unclear diagnostic statements.

Other frequently encountered areas of abstraction discrepancy for this measure set included data elements related to medications at patient arrival and discharge and contraindications to their use. Possible reasons cited for these discrepancies included abstractors' lack of familiarity with pertinent medication tables and data element definitions and inconsistency in the location of documentation of this information in the clinical record.

In some cases, hospitals made changes to the core measures medication tables or contraindication statements based on the judgment of clinical or medical staff, with abstraction consequently performed inaccurately. In one case, the hospital established and followed a clinical pathway for AMI patients that included contraindications to cardiac medications. However, the contraindications noted on the pathway did not mirror the technical specifications associated with the core measures. As a result, abstraction discrepancies occurred.

Heart Failure: The most frequent areas of disagreement between original abstraction and reabstraction for this measure set involved patient discharge instructions. Six distinct data elements comprise this performance measure, and the presence of all six, as well as an indication that the patient was given a written copy of the instructions, must be documented in the medical record.

Site visit findings indicated that generally, healthcare organizations had adopted distinct discharge instruction forms to capture these data elements. Quite often, however, all six required elements were not included on the form, or sections pertinent to these data elements were not completed.

Less frequently, the medical record lacked documentation that patients had been given a written copy of the instructions. Those hospitals that had documented the distribution to patients of pre-printed education pamphlets containing all the required data elements demonstrated the greatest degree of overall reliability for this measure.

Community-acquired Pneumonia: Abstraction discrepancies concerning this clinical measure set primarily centered around conflicting or inconsistent documentation of blood culture collection and antibiotic administration times in the medical record.

In many cases, disparate times were noted for the same occurrence in different sections of the medical record, leading to confusion as to which was the correct time. One hospital addressed this issue by designating a clock in the emergency department as the official clock to be used by all staff when noting germane times, such as blood culture draw time and antibiotic administration time, in the medical record. This resolved the issue for this hospital.

Pregnancy and Related Conditions: This performance measure set is comprised almost entirely of the general administrative data elements discussed earlier, and the points made in that discussion apply here. The most frequently encountered area of discrepancy between original abstractions and reabstractions for these data elements centered on admission type and source, for which medical record documentation appeared to be inconsistent with the data abstracted.

The hospitals that had the highest agreement rates for these measures were those in which core measures abstraction staff had communicated with registration staff to ensure a clear understanding of data definitions for these data elements, so that accurate assignment could be made at the time of patient admission.

Assess, Revamp, Educate, and More: Strategies of Successful Organizations

Clearly, although the general reliability of these data appears to be strong, opportunities for improvement of the reliability of core measures data abstraction still exist.

To a significant degree, the discrepancies between the original abstraction and reabstraction of the data elements can be attributed to medical record documentation issues. Much variation in data abstraction could be alleviated if healthcare organizations identified consistent record locations, formats, and documentation requirements for key data elements and then educated key clinical staff regarding this information.

The hospitals whose data abstraction demonstrated the greatest degree of reliability were those that had assessed their medical record forms and formats in light of core measure documentation requirements. When appropriate, they revamped these forms and formats to ensure that required data elements were documented accurately, consistently and, in some cases, using the appropriate verbiage.

Similarly, the organizations that undertook efforts to educate their physicians on the specifics of documentation requirements for such data elements as initial ECG interpretation for the AMI measure set or left ventricular systolic dysfunction (LVSD) for the AMI and HF sets demonstrated greater reliability than those that did not.

Healthcare organizations can take other actions to help ensure the reliability of the core measures data. For instance, they can educate and re-educate data abstraction staff as to the complete definitions of all data elements comprising the core performance measures. Literal, consistent interpretation of data definitions and data abstraction instructions is essential to achieving reliable data collection and abstraction performance over time.

For standardized measures, strict adherence to these definitions and instructions is not optional, regardless of occasional opinions to the contrary on the parts of clinical staff. Performance measurement systems are responsible for communicating current and accurate data element definitions and abstraction instructions to the hospitals with whom they contract for core measures support. Data element specifications have also been made available to hospitals online at www.jcaho.org/pms/core+measures.

Core measure data reliability would be further enhanced if all hospitals objectively assessed the organizational resources required to establish and maintain an adequate infrastructure to support core measures data collection and abstraction. Organizations should then work to ensure adequate allocation of those resources to these important organizational functions. During reabstraction site visits, it was observed that the data reliability of hospitals that had engaged adequate numbers of well-trained staff dedicated to completing these functions was superior to those that did not.

Finally, the reliability of the ORYX national core measures database could be significantly improved if all healthcare organizations established formal, structured performance assessment and improvement processes focusing on the core measures data collection and abstraction functions. Hospital staff interviews revealed that only half of the hospitals visited had established any processes to review the accuracy and completeness of core measures data abstraction. Most often, these efforts consisted of informal spot checking of data elements by peer or supervisory staff while conducting data entry or in response to questions brought forth by abstraction staff.

Far less frequently, more rigorous inter-rater reliability studies, consisting of reabstraction of a randomly selected sample of records by individuals other than the original abstractor, were performed on a scheduled basis. These findings are in interesting contrast to the fact that more than three-quarters of the hospitals visited during this study reported that they routinely and formally assessed their coding performance in a structured manner, generally through engagement of outside contractors.

Clearly, organizational concern for the accuracy and completeness of the core measures data collection and abstraction functions does not presently enjoy the degree of prominence attributed to the coding function. Considering the potential impact to healthcare organizations of core measures performance data, perhaps this estimation should be reassessed.

Note

1. Fleiss, J.L. *Statistical Methods for Rates and Proportions*, 2nd ed. New York: John Wiley and Sons, 1981.

Acknowledgment

Support for this project was provided by grant 1 U18 HS013728-01 from the US Department of Health and Human Service's Agency for Healthcare Research and Quality.

__Ann Watt__ (AWatt@jcaho.org) is associate project director, __Scott Williams__ is project director, __James Robertson__ is associate project director, __Richard G. Koss__ is director of department of health policy research, and __Jerod M. Loeb__ is executive vice president in the division of research with the Joint Commission on Accreditation of Healthcare Organizations in Oakbrook Terrace, IL. __Kwan Lee__ is director of health outcomes for Walgreens Health Initiatives.

Understanding the Core Measures

Core performance measures are a component of the Joint Commission's ORYX initiative, which is designed to integrate performance measurement into the accreditation process. The core performance measures were developed by panels of clinical experts and other key stakeholders to provide accredited organizations with objective feedback about their own performance to support internal performance improvement activities. They were also designed to be used externally to demonstrate accountability to patients, purchasers, and other stakeholders.

Core measures data are reported to the Joint Commission on a quarterly basis via performance measurement systems to which hospitals submit patient-level data. Performance measurement systems are Joint Commission-evaluated third-party vendors that have contracted with hospitals to accept and aggregate their core measures data in accordance with Joint Commission technical specifications.

To promote consistency of data abstraction over time within and across organizations, detailed data element definitions, abstraction instructions, and algorithms for computing measure rates were established for each core measure. These definitions and instructions have been incorporated by all performance measurement systems contracted to provide core measure data to the Joint Commission. The specifications are also available online at www.jcaho.org/pms/core+measures.

Article citation:
Watt, Ann, et al. "Keen Eye on Core Measures: Joint Commission Data Quality Study Offers Insights into Data Collection, Abstracting Processes." *Journal of AHIMA* 74, no.10 (November 2003): 20-25.