Methodology Matters—IX

Designing medical record abstraction forms

NAOMI J. BANKS

Center for Quality of Care Research and Education, Harvard School of Public Health, Boston, MA, USA

Data collection forms

The first step in clinical quality improvement frequently involves assessing the level of adherence to the clinical processes known or believed to contribute to improved patient health. Valid and reliable assessment of clinical processes can be conducted by measurement using an instrument composed of five elements. The first element is a clinical criterion which defines the recommended process of care. This is usually based on a guideline or protocol for the prevention or care of the condition under consideration. The second and third elements are the detailed identification of a population to whom the criterion applies and who are at risk for this condition, and the specific data elements required to determine patient eligibility and criterion conformance. The fourth element is the set of procedures for collecting data, and includes data collection instruments and data abstraction forms (the subject of this article). The fifth and final element is a performance measure containing specifications for data analysis which is used to calculate performance rates from the collected data.

Several types of paper instruments are used to collect data for performance measurement. Survey forms elicit facts and opinions directly from patients or their families. The respondent may complete a questionnaire or provide answers to questions asked by an interviewer. Survey instruments allow both patients and clinicians to furnish information about their opinions of the health care services and satisfaction with these services. In addition, patient reports can provide accurate data on the state of the patient's health and experiences with the health care they have received [1].

In contrast, data abstraction or data collection forms are used to abstract information from medical records. While medical record data have frequently been considered the 'gold standard' data (i.e. the most complete and reliable data available for understanding the patient care process) for quality reviews, they have inherent problems that should be recognized. With the exception of statements based on the clinician's own judgment of the patient's condition, medical record documentation is not primary data, i.e. from the patient him- or herself, but is filtered through the clinician, laboratory, or other source. In addition, since reviews are usually retrospective in relation to the provision of care, the reviewer cannot obtain clarification of missing or ambiguous data. Therefore, data collection instruments and rules for abstracting data must be carefully developed and highly explicit.

Coordination of data specification with the criterion development process

Data collection procedures are, by definition, an integral part of a performance measure [2]. When formulating clinical indicators and medical review criteria for quality of care evaluation, developers must specify the data which will provide evidence for the quality of care. Use of medical records to provide evidence for quality assessment raises the following questions:

- what types of data can provide proof that a criterion was met? Examples of these might be process data describing a clinician action or judgment, or outcome data describing patient health status before or after care;

- where can we find such documentation? Examples might be in inpatient, outpatient, or home care records;

- which parts of a record should or should not be used? Examples of these would be physician notes, nursing flowcharts, laboratory or radiology reports, discharge summaries;

- if two sources of documentation disagree, which takes priority?

- what are the various ways each concept can be expressed? For example, are the notations 'high blood pressure' (HBP), 'hypertension' (HTN), or specific numerical values for systolic blood pressure and diastolic blood pressure synonymous?

- what abbreviations and symbols are used for each concept? For example, 'normal', 'NL', 'WNL', '–', 'NAD'?

- how shall we interpret the lack of documentation for an indicator? as 'not done,' as 'normal,' or as 'can't evaluate'?

Address correspondence to Naomi J. Banks, Center for Quality of Care Research and Education, Harvard School of Public Health, 677 Huntington Avenue, Boston, MA 02115, USA. Tel: +1 617 432 1898. Fax: +1 617 432 3199.

Taking the time to specify data elements and instructions in this way is the critical first step to ensuring both the reliability and validity of performance measures constructed from medical record data.

Organization of the form

In general, the questions on an abstraction form should follow the order in which the information appears in the patient record or other source of data. If the data of interest are likely to be found in more than one place in a patient record or in more than one document, the questions should be grouped according to the source, or different forms should be provided for each source. For example, for the purpose of abstracting from the ambulatory record, the first section of the form is designed to collect information from all visit notes, and subsequent sections address data from laboratory tests, radiology reports, and consultation notes. Finally, the last section or a separate form can be used for data obtained from the patient's hospital record or discharge summary.

Data format and codes

A well-designed data abstraction instrument (Figure 1) is formatted to promote accuracy of data transcription, to limit the likelihood the abstractor will miss data items, and to promote, at a later stage, efficient and accurate data entry into the computerized database which will be used for the analysis. Transcription errors and data-entry errors can be reduced if the appropriate number of characters in each response field is indicated. The response fields (i.e. the spaces for entering data on the form) should be visually aligned and right-justified; no blank spaces (i.e. empty data fields) should be permitted. A leading zero is inserted when a single-digit number is abstracted into a two-character field.

As Figure 1 also shows, we recommend devising data codes that require few keystrokes. Number codes are preferred to letter codes because they permit more rapid data entry using the numeric key pad. Whenever possible, codes should have consistent meaning from question to question (e.g. 1 = yes, 2 = no), to enable the abstractor to memorize these codes. In cases where the desired data are missing from the record, the code '9,' or a combination of '9s', is commonly used rather than leaving the response field blank. Similarly, a specific code is selected for use in cases where a particular question is not applicable (the code '8' is commonly used for this purpose). When the responses on the form lead to branching, for example, 'If No, go to question 15,' the skipped questions are filled in either by the computer or by the data abstractor with the code for 'not applicable'. Finally, when a question on the form lists a selection of responses from which the abstractor must choose, but the list is not logically exhaustive, a code should be provided for 'other' to eliminate the possibility that the reviewer has to leave the space blank.

Form format and review design

Forms may vary in format. This will depend on whether, for the purpose of subsequent computer analysis, the data are recorded exactly as found in the medical record, or whether they must be interpreted by the abstractor. In the first case, a programmed computer algorithm applies the logic to determine whether a clinical criterion was met. In the second type of review, the criterion logic is applied by a highly-trained reviewer who answers a more complex question than those in the first type of abstraction form. The following example demonstrates these two methods.

For a clinical guideline that states 'If an abnormal hematocrit (HCT) is found, the HCT should be repeated within 3 weeks to confirm anemia' we could write a quality review

Data Item	Data Entry Field
Patient Identification	— — — — — —
Date of Birth (mm/dd/yy)	— — / — — / — —
Hematocrit done (1=Yes, 2=No)	—
Hematocrit result	— — · —
1. Was biopsy normal? 1=Yes 8=Not Applicable 2=No 9=Not Available	—
2. Type of referral 1=Surgery 7=Other 2=Gynecology 8=Not Applicable 3=Endocrine 9=Not Available	—

Figure 1 The features of an abstraction form demonstrate several techniques that help to ensure the quality of the abstracted data: use of a blank line for each character, alignment of data fields, avoidance of blank fields, consistency of meaning for codes, inclusion of a code for 'Other.'

criterion as 'For women over 50 years of age, if an initial HCT is <38.0, a second HCT is done <21 days'.

The abstraction form for the simple transcription method would require only the patient sex, age, HCT values, and the dates of all tests performed in a specified period of time. Little reviewer judgment is needed after the reviewer has been trained to locate the required data in the medical record. This method employs a computer program to evaluate the data and determine criterion conformance; the program applies a combination of decision rules for sex, age, HCT value, and dates.

Alternatively, the abstractor may be asked to provide a 'yes' or 'no' answer to the criterion question: if female >50 years, with a new HCT <38.0, was a second HCT done within 21 days?

Phrased this way, the question requires the abstractor, after considering all the available data, to use judgment to record in one data-entry field the correct answer derived from the combined answers to at least five questions:

● did the patient's age and sex qualify him or her for the criterion?

● was there a record of an abnormal HCT?

● if so, was it new?

● was another HCT done later, and if so, was the repeat HCT done on a date that is within 3 weeks of the initial abnormal HCT?

Assuming that the abstractor's interpretations and abstracting is correct, the answer that results from this series of questions and provides the coded entry in the data field is, in fact, the required indicator of conformance with the performance criterion, and no further analysis is necessary.

The second approach – rephrasing the quality criterion into a detailed review question – requires less start-up work. However, training for data collection is more difficult, and more sophisticated abstractors are needed, resulting in increased costs for data collection. Inevitably, because more judgment is required of the abstractors, reliability is more difficult to achieve.

The first approach, simple data transcription, requires greater start-up cost and involves programming a computerized algorithm to score each review criterion. However, once the computerized system is built, abstractors can review large numbers of cases, usually with higher degrees of accuracy, including those abstractors with less formal education. In practice, the review methods employed and the questions they generate usually fall somewhere between the two extremes in these examples.

Incorporating decision rules into data collection forms

For forms requiring a somewhat greater degree of abstractor judgment we have used a three-column format, as demonstrated in Figure 2. In addition to being organized chronologically by type and location of care, these forms have the following features:

● Column 1: sections and questions are numbered.

● Column 2: all permitted responses and their codes are listed; instructions for interpreting data are included; skip patterns are specified to permit abstractors to jump over questions that are not applicable.

● Column 3: variable names or codes are assigned to each data-entry field to help identify the data during analysis.

Abstraction procedures manual

A clearly written and generously illustrated manual of procedures should be written for the data abstractors' training and reference. The contents of this manual will include a restatement of the data elements, their location in the medical record, and their various synonyms, abbreviations, and interpretations as discussed above.

The abstraction manual may also contain examples of the appearance of the data sources the abstractor will use and examples of properly completed abstraction forms, including examples of situations that are commonly encountered in abstracting patient records, such as branching questions with skip patterns and missing or apparently conflicting data.

Pilot testing data forms and rules

A process of review and testing to ensure the reliability and validity of the data collection instruments is essential before implementing the quality measurement project. Investigators should examine drafts of the abstraction forms and procedure manuals for face and content validity[1] and determine whether all the data elements needed to evaluate criteria conformance have been included. An experienced reviewer will test the draft instruments by abstracting several patient records to ensure that the terminology and data formats of the forms are consistent with those found in the records.

The next step is to identify any ambiguous or confusing information which may have been incorporated during the drafting of the forms and manual. A determination is made whether the abstraction forms and procedures accommodate the broad variety of data configurations found in the patient records. For very extensive projects, where record reviews are conducted at many sites, 50 or more patient records should be abstracted during pilot testing in order to provide adequate representation of the different types of data that may be encountered in the main phase of the project. This field test of the performance measure is conducted by research personnel who have the same level of skills and experience

---

[1] A measure has face validity when it appears to measure what it is intended to measure, and it has content validity when all elements of the measure relate to the performance being measured and all relevant aspects of performance are covered by the measure.

A. Admission history and physical			
1. Was a risk assessment done preoperatively?	1 = Yes – if documentation of medical history AND current medications or note stating no medications taken 2 = No	RSKASS	——
2. Did the admission note include a history of organ failure, hypotension, GI bleeding (ulcer disease, gastritis, or other), or coagulopathy?	1 = Yes 2 = No	RSKBLD	——
3. Was the patient taking aspirin or an anticoagulant preoperatively?	1 = Yes 2 = No anticoagulants documented--GO TO QUESTION 5	ANCOAG	——
4. If QUESTION 3 = Yes, was anticoagulant/aspirin discontinued prior to surgery?	1 = Yes –≤24 hours prior to surgery 2 = Yes –>24 hours prior to surgery 3 = Not discontinued	DISCON	——
5. Was blood pressure determined preoperatively?	1 = Yes – if BP taken on or after admission but before surgery 2 = No	PROPBP	——
B. Intraoperative care			
1. Does the anesthesia flowsheet show monitoring of blood pressure every 5 minutes during surgery?	1 = Yes – every 5 minutes 2 = Monitoring at least 75% of the time at 5-minute intervals 3 = No – monitoring less than 75% of the time 7 = AA – if patient is on cardiopulmonary bypass machine 9 = UTD – anesthesia flowsheet not available	INTRBP	——
2. Was any NSAID given at any time on the day of surgery?	1 = Yes 2 = No	INSAID	——
C. Postoperative care			
1. Was the ventilator discontinued within 24 hours postoperatively?	1 = Yes 2 = No 7 = AA – if physician note states attempt to wean patient or reason ventilator needed 8 = NA – patient not on ventilator	DCVENT	——
2. Was the patient on a ventilator for more than 48 hours?	1 = Yes 2 = No	POVENT	——
3. Was blood pressure taken at least twice daily for the first 10 days postoperatively?	1 = Yes 2 = No	PSTBP1	——
4. While on ventilator or if Question A2=Yes, was patient given peptic ulcer prophylaxis at least once daily?	1 = Yes – if daily H_2 blockers, antacids, Carafate, or Omeprazole given p.o., by IV, or by NG tube 2 = Not given or not given daily 7 = AA – patient refused medication 8 = NA – Question A2=No (no risk factors) and patient never on ventilator	ULCRRX	——
5. Were NSAIDs given prior to postoperative day 4?	1 = Yes – if any NSAIDs given before post-op day 4 2 = No - or given on or after post-op day 4 7 = AA – the only NSAID taken before post-op day 4 was ≤325 mg of aspirin	PONSAI	——

AA = Acceptable Alternative
UTD = Unable to Determine

Figure 2 This figure shows portions of a data abstraction form for a number of clinical care processes experienced by patients who developed postoperative gastrointestinal bleeding. Illustrated in this figure is a three-column format recommended to improve the quality of data abstracted from medical records and to facilitate data entry and analysis. Instructions for abstractors in column two regarding decision rules for coding enhance the reliability and validity of the data.

as those who will conduct the large-scale review. Intra-rater reliability is determined when an experienced abstractor uses the abstraction forms twice to abstract from the same batch of records and the two sets of results are then compared. Inter-rater reliability involves a comparison of the results from two abstractors of equal skill, each using the same forms once. By calculating the degree of match between the sets of coded information from the two abstractors, the project's directors can obtain a measure of inter-rater reliability. Both methods are employed to determine how much of the variation between the sets of data collected is due to the abstraction instruments themselves.

The data abstracted during these two types of reliability tests are compared for agreement by using standard statistical packages. These calculate the percentage of agreement for each data item reviewed and provide a list of those items that disagree. By determining the items which have high rates of disagreement, we can identify the data elements that reduce the reliability of the instrument. Agreement, and therefore reliability, may then be improved in one or more ways: by providing more explicit decision rules in the instruction manual, by better training of the abstractors, or by adjusting the design of the abstraction forms. Re-review and revision of the performance measure is repeated until the agreement rate reaches an acceptable level, for example, 95% agreement.

Clinical performance measurement depends on the thoughtful contributions of clinicians, medical records specialists, and health services researchers. Ensuring data integrity is perhaps one of the most challenging tasks in the planning and implementation of a measurement project. Data abstraction forms designed and tested with the methods described in this article will contribute greatly to the reliability and validity of research findings.

References

1. Hargraves JL, Palmer RH, Zapka J *et al.* Using patient reports to measure health care system performance. *Clin Perform Qual Health Care* 1993; **1**: 208–213.

2. Agency for Health Care Policy and Research. *Using Clinical Practice Guidelines to Evaluate Quality Of Care.* Vol. 2, *Methods.* Washington, DC: Department of Health and Human Services, 1995.