Resolving Confidentiality Barriers in Research Data Collection

by Shannon H.S. Houser, MPH, RHIA, Molly Engle, PhD, Joseph E. Schumacher, PhD, Ellen Caldwell, PhD, Connie L. Kohler, PhD, Sharon T. Phelan, MD, James M. Raczynski, PhD, and Kim D. Reynolds, PhD

How can HIM professionals ensure data collection consistency without access to earlier records? When protecting patient anonymity posed a challenge in a research study, HIM professionals took a team approach to abstracting records that maintained quality without compromising confidentiality.

Numerous research studies are conducted using medical record reviews, and most include an identification variable enabling data collectors to periodically conduct quality control measures on previously collected data. In a recent research project conducted by the University of Alabama at Birmingham's Behavioral Medicine Unit, Division of Preventive Medicine, quality control through repeated measurement of specific medical records was impossible, due to the sensitivity of the information collected. Ensuring data reliability without employing traditional repeated measurements with patient identifiers became a primary concern in the project design.

Accepting the Challenge

The four-year study, funded by the National Institute on Drug Abuse, developed, implemented, and evaluated a clinical behavior-change training program for obstetricians and obstetrical staff from 10 of 27 available community-based obstetric practices. The program's goal was to increase the assessment, management, and referral of pregnant women who used or abused alcohol, tobacco, and other substances. The program used a variety of intervention methods and medical record reviews of target behaviors were conducted at each practice at three time points. Data collected from medical records focused on assessment of behaviors related to past and current use of tobacco, alcohol, and other drugs; active or prior obstetric, gynecologic, social, psychiatric, or medical problems and trauma; and any substance abuse management/counseling, or referral activities conducted by healthcare providers. Because of the sensitivity of such information, the Institutional Review Board (IRB) disallowed the collection of demographic data (names, dates of birth or identification codes) that could be used to identify individual patients, preventing data collectors from accessing the records for data verification and error correction. Thus, a strategy to assure data reliability while protecting patient anonymity and confidentiality emerged as an obstacle in the study.

All data collected were aggregated and analyzed as group data. No individual professional or patient was identified. A total of 2,226 records were reviewed at 10 practices across the three data collection time frames.

Sticking to the Protocol

Before data could be collected, two major issues related to data collection had to be resolved. Because of the IRB directive, once the initial review of a patient's record was completed, that record could not be later identified for an additional review for data quality verification. Thus, reliability of data collection became a critical issue. Further, deviation from the data collection protocol posed threats to the validity of the data collected. Generally, when HIM professionals review records, they make a concerted effort to glean as much data or information from the records as possible. In a non-restricted study, an experienced abstractor (often a coder trained to abstract for reimbursement) may collect data not specified by the protocol to make additional information available to the researchers. This often leads to an extension of the research objective and the review may be repeated to collect additional data from all records related to emergent variables. However, in this study, data collection had to be conducted strictly within the protocol design and the IRB restrictions, and any interpretation or deviance by the abstractors was not allowed.

Plotting the Course

Investigators developed a six-step data collection process that began with the assessment of the project's needs and identification of problems related to reliability and validity and ended with quality control mechanisms (see table on page 64). Development of the medical record review protocol, forms design, training, pilot testing and form revision, and actual data collection completed the process.

Because reliability of the collected data was a key factor in this research, the investigators were challenged to develop a method to ensure reliability because retrospective controls could not be applied. First, two abstractors, rather than one, were employed. After training in the abstraction methodology, they worked as a team at each data collection site. Each abstractor reviewed each record independently and after completing a data collection form for one medical record, they exchanged records and independently abstracted the new record. They then compared and discussed their individual reviews of each record and resolved any differences based on the objectives of the research before progressing to a new record. Thus, there was a 100 percent reliability check by two reviewers on all the records.

Further, the abstractors were blinded from intervention group assignment. At the time of the records review, the abstractors were not informed whether the patient was in the group that first received the intervention or the group that received the intervention second, as all practices received the intervention in a delayed sequence. If they had known that the patient had received intervention, the abstractors might have expected, and thus found, evidence in the records indicating progress and inadvertently introduce a bias into the data.

Road Test

Before data collection started, abstractors at an OB/GYN clinic outside in the study group conducted a pilot test of the record abstraction form and the review process and procedures. As a result of the pilot test, problems associated with information collection were identified and the abstraction form was revised. This pilot test assured that the form design was appropriate and that the collected data would meet study requirements. Also, it was a useful way to train the abstractors without contaminating the study practices.

The record review pilot test also produced an estimate of the average time required to review a record, which was greater than expected, due partly to the process of resolving differences between abstractors. Reviewing paired records took more time when multiple differences occurred and each difference had to be referenced to the original record and agreement negotiated by the abstractors. The estimated average time varied among practices depending on the complexity of the practices' records. This time was lessened once abstractors became familiar with the records at each practice and fewer points of disagreement occurred.

Maintaining Skills and Standards

Both abstractors were credentialed HIM professionals with prior experience and trained by project investigators in the use of study protocol and data collection instructions. Continued training was performed periodically until the abstractors could demonstrate at least a 90 percent agreement on data collected from a sample of records. Because intervals between data collection periods were six months, project investigators retrained the abstractors before each additional data collection period to ensure consistent data collection. Employing two HIM professionals, providing continuing training, and retaining the same abstractors throughout the project insured reliability, validity, and quality of data collected.

Because record reviews were conducted at multiple practice sites, documentation and patient record forms and their use differed among practices. For quality control purposes, the project data manager went to the first selected practice at the beginning of each data collection period, randomly selected two records for review, and then compared the results with the findings of the abstractors. Any questions the abstractors had about a particular practice's documentation and forms were resolved at the beginning of the data collection. If abstractors were unclear about how data should be coded at a practice, they contacted the project data manager first and tried to resolve the problems on site. If additional questions arose, abstractors recorded the data in a different-colored ink, and these problems were resolved at the weekly debriefing.

Avoiding Obstacles

The data collected from record reviews was designed to provide evidence of changes in health professionals' behavioral patterns in addressing assessment, management, and referral of pregnant women who used or abused substances. Without the steps taken to ensure quality control and the reliability of collected data, the results could have been questionable. The major hurdles to navigate were:

- Data consistency. While a high degree of reliability may be achieved by employing more than one data collector in a study using anonymous medical records, there is a need to assure standardized data collection by using the same well-trained abstractors throughout the project.

Data consistency could be questioned if different abstractors were used for different data collection periods without a quality control check. In most studies of this type, only a sample of abstracted records are reviewed for reliability. In this study, 100 percent of all records abstracted in the study were reviewed twice.

- Cost and expense. Each record was reviewed twice, and the abstractors compared their findings after the reviews and resolved any differences. This method entailed more than twice the time a single review would, and it also cost more than twice as much because of the need for two salaries paid for longer periods of time. There were also additional costs for travel reimbursements.
- Professional compatibility. It was essential that the two data collectors be compatible, independent, and consider each other as equals. Recruit data collectors with similar professional and educational backgrounds to avoid potential setbacks.
- Validity of reported data. The study only measured what was documented in the medical records by the healthcare practitioners. For outcomes not documented in the records, no measurement could be made. Intervention methods, like patient advisement by healthcare professionals and the presence of educational posters provided by the project as a part of patient education, were not documented in the medical records. Thus, these data could not be quantified or measured from the records documentation as a part of the intervention results. To obtain a more accurate assessment of interventions, an additional measurement, such as a survey or interview script of office staff, could be designed to report these other data.

The dual approach to record review proved to be very effective in assuring data reliability when patient identifiers were unavailable. The extensive project-specific training and retraining for the abstractors used throughout the project served as a data control mechanism to support the findings of the project. The use of two unbiased abstractors, both of whom were professionally educated in HIM, expedited the process.

This method could be used in other projects in which patient confidentiality is paramount and individual medical record identification is restricted, such as in HIV/AIDS research. This approach to medical records abstracting makes research possible in highly sensitive areas that otherwise might be avoided because of methodology constraints, many of which are most in need of research. The education and skills of HIM professionals permit them to make a significant contribution to this type of research study.

Six-Step Data Collection Process:

1. Needs assessment/problem identification
2. Development of medical record review protocol and form design
3. Training abstractors
4. Pilot testing/form revising
5. Data collection
6. Quality control

Acknowledgments

Special thanks to Donna Slovensky, PhD, RHIA, and Howard Houser, PhD, School of Health Related Professions; Pamela Mott, Department of English at University of Alabama at Birmingham; Kay Clements, RHIA; and Barbara Yarber, RHIT.

This research was made possible, in part, with funds from the National Institute on Drug Abuse and the Division of Preventive Medicine, Behavioral Medicine Unit, University of Alabama at Birmingham.

Shannon H.S. Houser is a clinical data analyst in the School of Medicine at the University of Alabama. She can be reached at shouser@dopm.uab.edu. Joseph E. Schumacher and Sharon T. Phelan are associate professors in the School of Medicine at UAB. Connie L. Kohler is an assistant professor and Kim D. Reynolds is a scientist in the School of Public Health at UAB. James M. Raczynski is a professor in the Schools of Medicine and Public Health at UAB. Molly Engle is Associate Professor at Oregon State University. Ellen Caldwell is a biostatistician at the University of Washington.

Article citation:
Houser, Shannon, et al. "Resolving Confidentiality Barriers in Research Data Collection." *Journal of AHIMA* 71, no.3 (2000): 62-64.